CELLULAR IMMUNOLOGY 152, 556–568 (1993)

I. A Novel Monoclonal Human IgM Autoantibody which Binds Recombinant Human and Mouse Tumor Necrosis Factor-α

PETRA BOYLE, KENNETH J. LEMBACH, AND GAYLE D. WETZEL

Pharmaceutical Division, Biological Products, Cellular and Molecular Biology Research, Miles Inc., P.O. Box 1986, Berkeley, California 94701

Received May 28, 1993; accepted August 6, 1993

Several monoclonal human IgM antibodies to recombinant human tumor necrosis factor-α (rhTNFα) have been generated and partially characterized. The F78-1A10-B5 monoclonal antibody (mAb) (B5) binds to rhTNFα with a titer comparable to three high-affinity neutralizing mouse mAbs, when tested by ELISA. However, the B5 mAb binds relatively weakly to soluble rhTNFα. It appears to bind to epitopes on rhTNFα distinct from those bound by the mouse mAbs for three reasons. First, preincubation of plate-bound rhTNFα with mouse mAbs does not decrease or compete subsequent B5 mAb binding. Second, rhTNFα complexed to the mouse mAbs can still be bound by B5 mAb. Third, the mouse mAbs neutralize TNFα cytotoxicity whereas the B5 mAb does not. Binding analyses indicate that this human IgM autoantibody binds to both human and mouse recombinant TNFα, but not to other antigens commonly recognized by polyreactive natural IgM autoantibodies. The high level of amino acid identity between the human and mouse TNFα molecules suggest that the B5 mAb is monospecific for a given epitope shared by these two forms of TNFα. This spectrum of characteristics makes B5 a novel mAb.

INTRODUCTION

Tumor necrosis factor-α (TNFα) is a pluripotent and pleiotropic cytokine (for recent reviews see Refs. 1–5). Although produced principally by activated macrophages, its synthesis and secretion have also been observed using granulocytes (6), tonsil B cells and B cell lines (7), natural killer cells (8), T cell lines, and peripheral blood T cells (9). TNFα has been reported to induce thymocyte growth (10), B cell growth and maturation (11, 12), hemorrhagic necrosis *in vivo* (13), weight loss, cardiovascular collapse, and multiple organ failure (1, 2). Naturally, these latter activities are the source of the clinical interest in TNFα.

During septic shock, as well as inflammatory diseases, synthesis and secretion of TNFα, IL-1, IL-6, and IL-8 have been documented. Hence, the immune systems of some individuals are exposed chronically to these cytokines. Indeed, low-affinity antibodies to TNFα (14, 15) and to IL-1α (15) have been reported. These anti-TNFα autoantibodies may, however, not be specific (16). One peculiar feature of human serum, as well as sera from other animals, is its content of natural and so-called polyreactive antibodies. These are usually IgM antibodies which bind to various autoantigens with low affinity (17, 18). Hence the autoantibody-like reactivity to human TNFα might be expected to be low affinity and probably cross-reactive with several

0008-8749/93 $5.00

other antigens. Despite these considerations, some high-affinity neutralizing antibodies to IL-1α in normal (19) or patient (20) sera have been reported.

We have produced several human monoclonal antibodies (mAbs) which bind to human TNFα. Here we characterize some of these IgM mAbs. Although many showed either intermediate or broad polyreactivity, as defined by binding to a panel of antigens, one particular mAb, B5, is characterized in detail since it appears to be monospecific for a determinant shared by mouse and human TNFα.

MATERIALS AND METHODS

Reagents

P. Stadler (Bayer, Wuppertal, Germany) provided recombinant human TNFα (rhTNFα). The recombinant mouse TNFα (rhTNFα) and recombinant human lymphotoxin (rhTNFβ) were purchased from Genzyme. Human IgG Fc fragments were purchased from Chemicon. Collagen was purchased from Boehringer Mannheim. All other antigens were purchased from Sigma. TNF was biotinylated using standard techniques; briefly, *N*-hydroxysuccinimidyl ester of biotin was added to TNF dissolved in 50 m*M* $NaHCO_3$, pH 8.5, for 15 min, quenched with NH_4Cl, and then dialyzed to remove unreacted biotin. The A10G10 anti-hTNFα IgG_1 mouse mAb was generated in collaboration with Chiron Corporation (Emeryville, CA). The A6 and B6 mouse IgG_1 mAbs were generated from mice hyperimmunized in our laboratory. All three mouse mAbs neutralize TNF cytotoxicity and have been described (21). These mAbs were purified by affinity chromatography. The polyreactive IgM mAbs 1A6B5F and F2.2.34 were produced and characterized elsewhere (22). The 7T1 human IgM mAb was produced and provided in ascites form by Professor M. Hoffmann (New York Medical College, Valhalla, NY). The 6F11-E4 (6F11) human IgM mAb is specific for Fisher type 2 *Pseudomonas* LPS and was produced by an EBV-transformed B cell lymphoblastoid line obtained from Genetic Systems Corporation (Seattle, WA). It serves as an isotype-matched control mAb for the human anti-rhTNFα mAbs. The C7F7 mAb is a mouse IgG_1 anti-hFVIII developed in collaboration with Genentech Inc. (South San Francisco, CA) and is used as an isotype-matched control mAb for the mouse anti-rhTNFα mAbs. Biotinylated mouse anti-human IgM, goat anti-mouse IgG, and biotinylated goat anti-human IgG were purchased from Jackson Labs. Biotinylated goat anti-mouse IgG was purchased from Zymed. Avidin-coupled HRP and avidin-coupled alkaline phosphatase were purchased from Zymed.

ELISAs

Antigens or capture antibodies (anti-immunoglobulin antibodies) were coated to plastic plates in carbonate/bicarbonate buffer, or PBS containing 20 μg/ml BSA, overnight at 4°C or for 3 hr at 37°C. Secondary incubations were carried out overnight at 4°C or at room temperature for a period of 2 hr or less. Secondary antibodies were biotinylated and their binding was detected using avidin-coupled HRP and avidin-coupled alkaline phosphatase.

Hybridoma Production

The human IgM mAbs listed in Table 1 were produced by fusion with the mouse P3X63Ag8.653 nonsecreting myeloma. Peripheral blood mononuclear cells from a

CMV-positive donor were separated by centrifugation on Ficoll, treated with L-leucyl leucine methyl ester, incubated *in vitro* with antigen, and subsequently transformed with EBV. Transformants were distributed at limiting concentrations and cells producing antibody binding to TNF were fused and subsequently subcloned. The B5 hybridoma was subcloned a minimum of five times. The H5 and 7T1 mAbs were produced by fusion of human tonsillar cells immunized *in vitro* by Dr. M. Hoffmann. Monoclonal human IgM antibodies were affinity purified by standard techniques, as described by others (23) for use in subsequent experiments.

Cytotoxicity Assay

To assess the TNF-neutralizing ability of various mAbs, the assay described by others (21) was used with the following minor modifications. Briefly, 20 pg/ml TNF was incubated overnight with 60,000 WEHI 164 cells and the test mAb. Viable cells were then detected by crystal violet staining and reading optical density at 570 nm.

Western Blotting

Recombinant hTNFα (100 μg/ml plus 100 μg/ml BSA) and recombinant mTNFα (5 μg/ml with 500 μg/ml BSA) were electrophoresed in the presence of β-mercaptoethanol and SDS on 12% polyacrylamide gels. Proteins were then electro-transferred to nitrocellulose paper which was then blocked with BSA. Test mAbs were allowed to bind and were subsequently detected with biotinylated anti-immunoglobulin reagents. Streptavidin–HRP was then added followed by substrate.

RESULTS

The Monoclonal Human IgM Antibody B5 Binds to Solid-Phase Recombinant Human TNF

Several hybridomas secreting monoclonal anti-rhTNFα antibodies have been established in our laboratory. Table 1 presents an endpoint titer analysis comparing a

TABLE 1

Comparison of Solid-Phase ELISA Format rhTNFα Binding by Several Monoclonal Antibodies

mAb	Endpoint titer (ng/ml)	Ig class
A10G10	0.6	Mouse IgG
A6	0.15	Mouse IgG
B6	0.08	Mouse IgG
F78-1A10-A1	0.3	Human IgM
F78-1A10-B5	0.6	Human IgM
F80-1B9-F12	0.15	Human IgM
F81-4E3-D6	9.8	Human IgM
F83-1D6-B6	625.0	Human IgM
D83-1D6-F6	1250.0	Human IgM

Note. ELISA plates were coated with 2 μg/ml rhTNFα. The indicated mAbs were added in titrated concentrations and binding was assessed spectrophotometrically using enzyme-linked antibodies. The minimum mAb concentrations yielding detectable rhTNFα binding are shown.

panel of six human IgM mAbs to three high-affinity neutralizing mouse mAbs, A10G10, A6, and B6. Soluble rhTNFα was bound to plastic plates and each mAb was added in titrated doses. The final concentrations of mAb at which rhTNFα binding was observed are presented. B5 and F12 (F80-1B9-F12) and A1 (F78-1A10-A1) were three of the best human IgM mAbs by this analysis, showing endpoint titers in the subnanogram/milliliter range, comparable to the three mouse mAbs.

Figure 1 presents a more extensive comparison of the human B5 and mouse A10G10 mAbs. Binding of both mAbs was concentration dependent regardless of TNF coating concentration. The B5 mAb appeared to bind slightly better than A10G10 with high TNF coating concentrations. As the TNF coating concentration was reduced, however, the binding of B5 decreased more rapidly than that of A10G10. This is consistent with B5 having a lower affinity than A10G10 for rhTNFα.

B5 mAb Binds to an Epitope on rhTNFα Different than Those Bound by Three Mouse Anti-TNF mAbs

Competitive binding experiments have shown that A10G10 and B6 recognize similar epitopes on rhTNFα whereas A6 recognizes a different epitope (data not shown). To examine the epitope binding specificity of B5, competitive binding experiments were performed using the mouse mAbs and B5. The mouse mAbs were added at different concentrations to ELISA plates previously coated with TNFα. An optimum concentration of B5 mAb was then added and binding was subsequently detected with biotinylated anti-human IgM. If the mouse mAbs recognize the same epitope as B5 mAb, they should inhibit B5 mAb binding in a concentration-dependent manner. As shown in Fig. 2A, the binding of the mouse mAbs to plate-bound rhTNFα is concentration dependent. Figure 2B shows that none of the mouse mAbs interfered with the rhTNFα binding by a fixed amount of B5 mAb, even at concentrations of the mouse mAbs significantly in excess of those required for maximal binding to the plate. These data suggest that B5 recognizes an epitope on rhTNFα different from those recognized by A10G10, A6, and B6.

To support this finding, rhTNFα was added to ELISA plates previously coated with the combination of A10G10, B6, and A6 mAbs. B5 mAb was then added to test whether it could bind to rhTNFα complexed to, or captured by, the mouse mAbs. Figure 3 shows that B5 and all the other human IgM mAbs, except 7T1, bound to rhTNFα complexed to mouse mAbs. Binding of the human mAbs was not seen in the absence of rhTNFα, demonstrating specificity for some component of rhTNFα. The failure of 7T1 mAb to bind to complexed TNF may be simply due to a low affinity. These results support the conclusion that the human IgM mAbs B5, F12, A1, B6, and D6 and the three mouse mAbs recognize different epitopes on rhTNFα.

B5 mAb Is Not Polyreactive

Since B5 mAb is a human IgM which binds to human TNFα, and therefore has properties which define it as an autoantibody, it was important to determine the specificity of this mAb and assess its polyreactivity. We chose a panel of human and nonhuman antigens typically used to define polyreactivity. Binding to these antigens by B5 mAb, A10G10, two control polyreactive human IgM mAbs 1A6B5F and F2.2.34 (22) and two other human IgM anti-TNF mAbs were compared. The results have been normalized for each antibody to allow direct comparison. Figure 4 presents the

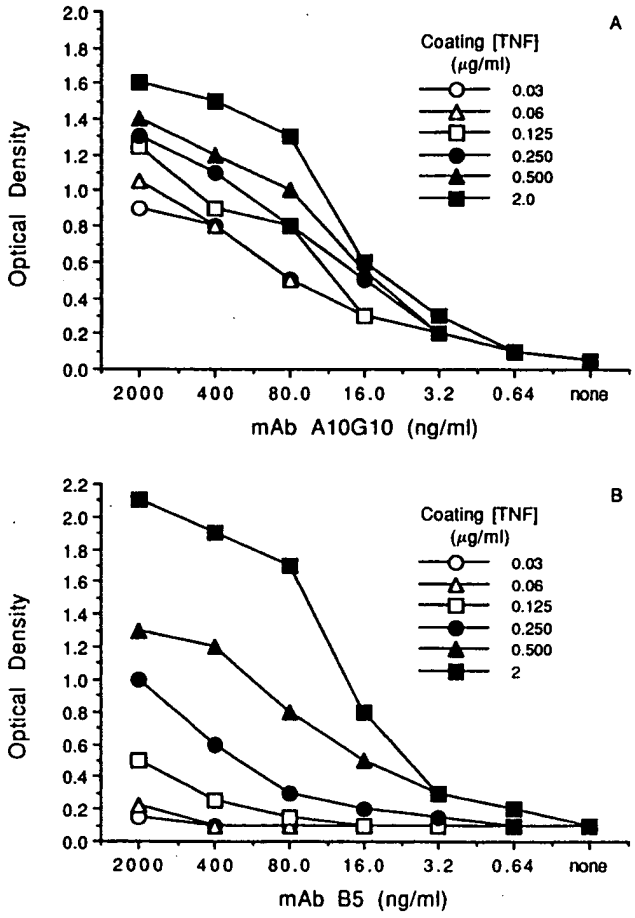


FIG. 1. Comparison of solid-phase ELISA format binding to rhTNFα of human mAb B5 and mouse mAb A10G10. ELISA plates were coated with various concentrations of TNF and titrated doses of mAb were then allowed to bind. Shown are the binding curves for each antibody for the various TNF coating concentrations.

data from one of four similar experiments. The mouse mAb A10G10 binds specifically to rhTNFα and none of the other antigens. In contrast, the polyreactive mAb 1A6B5F binds to virtually all of the antigens tested. The same was true for the other polyreactive mAb F2.2.34 although binding to BSA and TNF was much stronger than that seen with the other antigens. The B5 mAb showed specificity for rhTNFα. No binding by B5 mAb to recombinant human lymphotoxin (rhTNFβ) nor to any of the other antigens tested was observed. These data provide evidence that the B5 mAb is not polyreactive.

In contrast, the 7T1 and H5 human IgM mAbs bind to human Fc fragments, indicating a rheumatoid factor nature. These two antibodies also bind to insulin and 7T1 binds BSA, as well. The control polyreactive mAbs appear to define two classes of polyreactivity, one being very broad in specificity and the other being more restricted in the antigens recognized. The 7T1 and H5 mAbs belong to the more restricted class of polyreactive mAbs. The F12 anti-TNF mAb binds to human TNFα but only marginally to other antigens.

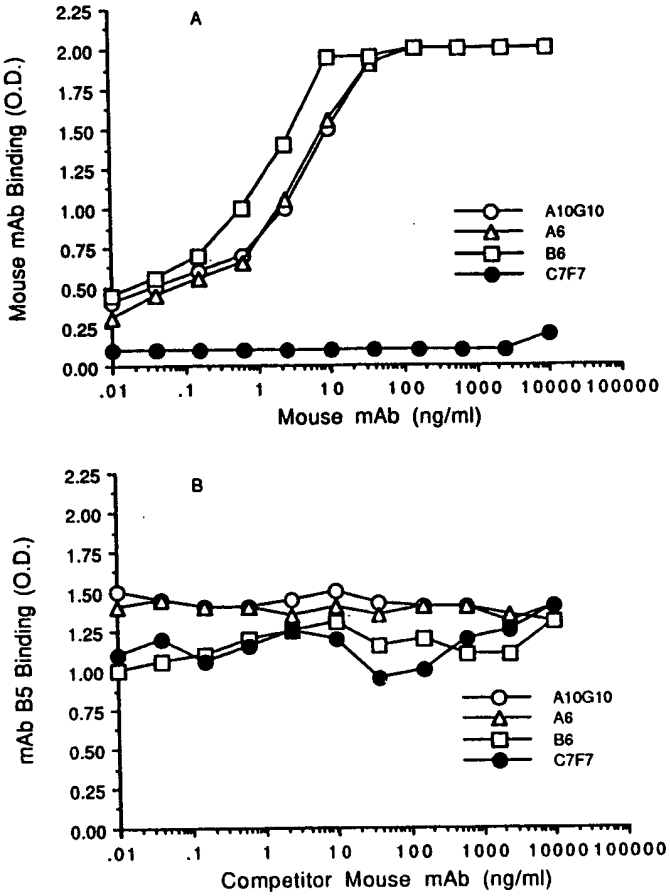


FIG. 2. Lack of TNFα binding competition between mouse mAbs and B5 mAb. Panel A shows three mouse anti-TNF mAbs and control C7F7 anti-rFVIII mAb binding to rhTNFα. Panel B shows the lack of inhibition of the binding of an optimal concentration of B5 to plate-bound TNF when the indicated concentrations of mouse mAbs were first allowed to bind.

B5 mAb Binds to Recombinant Mouse TNFα

During the course of analyzing the specificity of the B5 mAb, we noticed that it also bound to mouse TNFα. To demonstrate this, we first captured mouse TNFα with a neutralizing hamster monoclonal antibody and then allowed B5 to bind to this complex. Figure 5 shows the results of this kind of experiment. The binding of B5 was dependent on both the concentration of B5 present and on the concentration of hamster antibody used to coat the plates. No binding was observed when mouse TNFα was not added, indicating the specificity of B5 binding in this system. Other experiments not shown revealed binding to mouse TNFα by the F12 mAb.

B5 mAb Binds to Soluble rhTNFα with Detectable but Low Affinity

Next, we assessed the ability of the B5 mAb to bind to soluble rhTNFα. ELISA plates were coated with anti-human IgM and B5 was then added. The ability of the

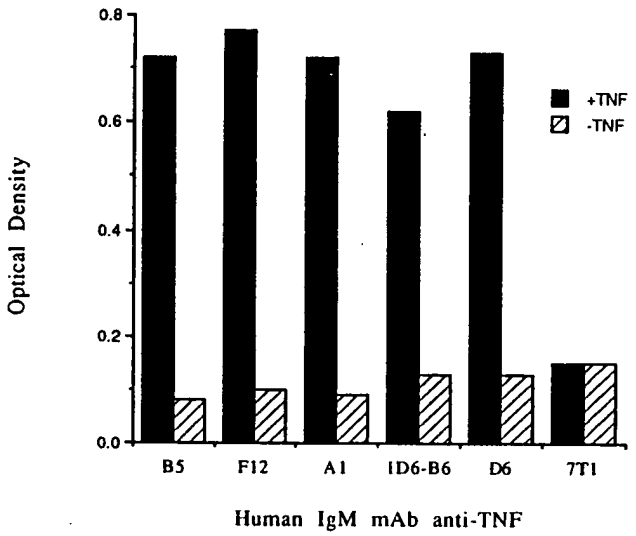


FIG. 3. Binding of human IgM anti-TNF mAbs to rhTNFα captured and presented as a complex by the combination of plate-bound mouse mAbs A10G10, B6, and A6. ELISA plates were coated with the three mouse mAbs and then incubated with rhTNFα. Plates were washed and 20 μg/ml of the indicated human IgM mAbs was then allowed to bind. Solid bars show the binding of the human IgM mAbs to the three mouse mAbs which had been incubated with TNF, and the hatched bars show binding of the IgM mAbs when the attached mouse mAbs had not been exposed to TNF.

bound B5 mAb to capture biotinylated rhTNFα was then determined. Figure 6 compares the binding of soluble TNFα to A10G10 and B5 under these conditions. Although both mAbs were found to bind soluble rhTNFα, the concentration of B5 mAb required for equivalent binding of the antigen was at least 300-fold greater than that of A10G10. Furthermore, binding of soluble TNFα to immobilized B5 did not saturate with the concentrations of B5 tested. These results are consistent with a low-affinity binding of rhTNFα by B5 mAb. Indeed, attempts to measure the binding constant of B5 mAb revealed an affinity too low to calculate by conventional methods (data not shown).

Soluble rhTNFα binding by B5 was also demonstrated by coating plates with anti-IgM, capturing B5, and then adding unmodified soluble rhTNFα. A10G10 was added next and its binding to this B5-complexed form of rhTNFα was detected with biotinylated anti-mouse IgG. Figure 7 compares the abilities of B5 and a control human IgM, 6F11, to capture and present soluble rhTNFα to A10G10. Although some nonspecific binding was seen with the control mAb, B5 mAb bound approximately four- to eightfold more rhTNFα in this experiment. These data are consistent with a low binding constant of B5 and add further support for the concept that B5 mAb and A10G10 mAb recognize different epitopes on rhTNFα.

B5 mAb Recognizes rhTNFα in Western Blots

Figure 8 shows the results of an experiment using Western blotting to demonstrate B5 binding to denatured TNFα. In lanes A–F, binding to mouse TNFα was examined and in lanes G and H binding to human TNFα was examined. The 6F11 antibody did not bind to either TNFα species and so provides a specificity control. All of the

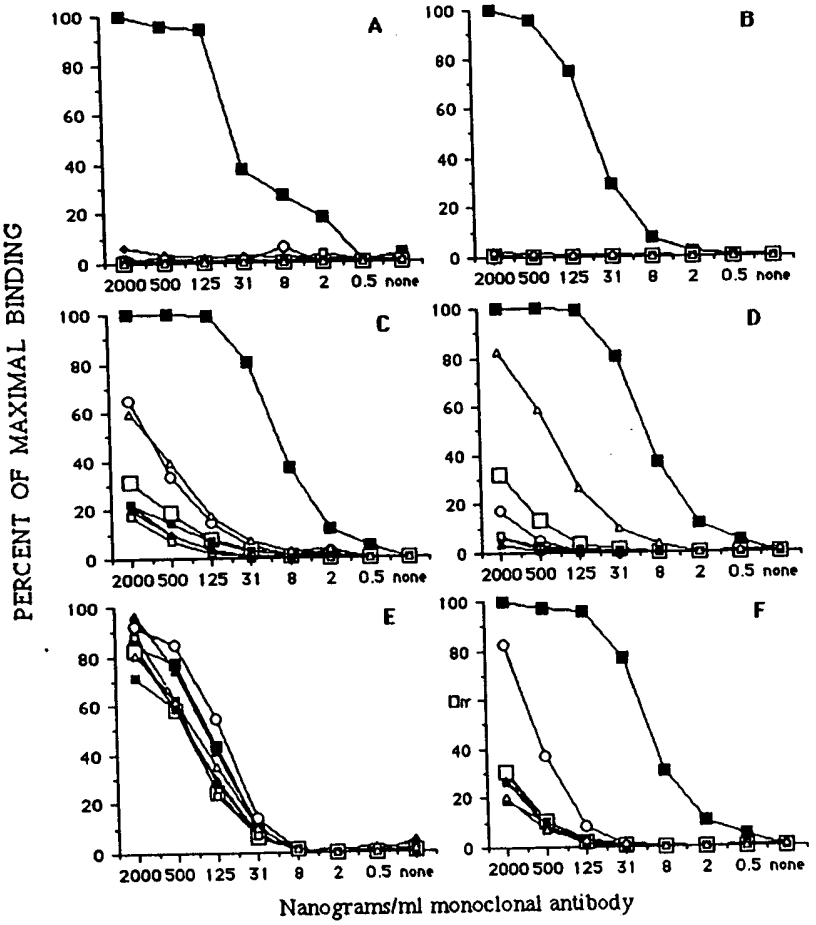


FIG. 4. Analysis of the binding specificities of several mAbs. Plates were precoated with either recombinant human TNFα (■), recombinant human lymphotoxin (♦), human insulin (□), porcine thyroglobulin (▲), BSA (○), ssDNA (●), dsDNA (□) or human IgG Fc fragments (△). Mouse mAb A10G10 is shown in A. Human IgM mAbs B5, 7T1, H5, 1A6B5F, and F2.2.34 are shown in B, C, D, E, and F, respectively.

human IgM mAbs, 7T1, H5, 1A6B5F, and B5, bind to mouse TNFα. Furthermore, the B5 antibody also binds to human TNFα, under these conditions. The polyreactive mAbs bind human TNFα in Western blots as well (data not shown). These results suggest that B5 may recognize a linear epitope of rhTNFα.

B5 mAb Does Not Neutralize the Cytotoxicity of rhTNFα

The TNF-sensitive cell line WEHI 164 was used to assess the ability of B5 mAb to neutralize TNFα cytotoxicity. Figure 9 shows that A10G10 clearly neutralizes rhTNFα in a dose-dependent manner, as previously demonstrated by others (21). At no concentration of B5, however, was any neutralization of rhTNFα observed. The same is true for the three other human IgM anti-TNFα mAbs, B6, F12, and 7T1. These data add further support to the idea that B5 and A10G10 bind different epitopes of TNF and are consistent with the weak ability of B5 mAb to bind soluble rhTNFα.

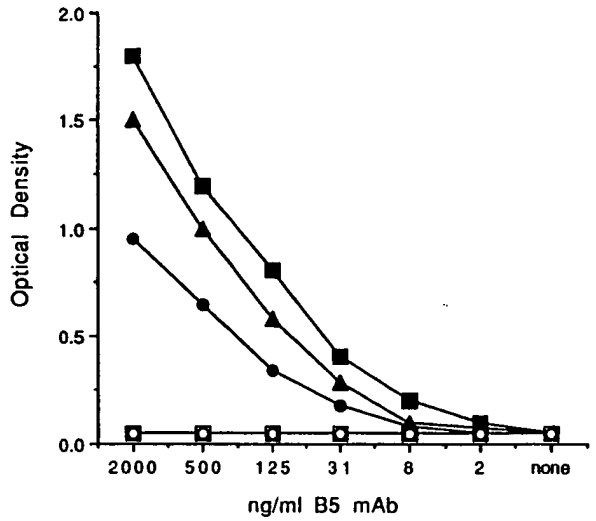


FIG. 5. Binding of B5 to recombinant mouse TNFα. Plastic plates were precoated with a neutralizing monoclonal hamster anti-mouse TNFα antibody at 8 μg/ml (squares), 4 μg/ml (triangles), and 2 μg/ml (circles). Recombinant mouse TNFα was then added at 2 μg/ml (filled symbols) or was not added (open symbols). Human mAb B5 was then allowed to bind at the concentrations indicated. Binding was then assayed with anti-human IgM antibody.

DISCUSSION

To our knowledge, this is the first report of a monoclonal human autoantibody specific for human and mouse TNFα. It is unclear whether the CMV-seropositive

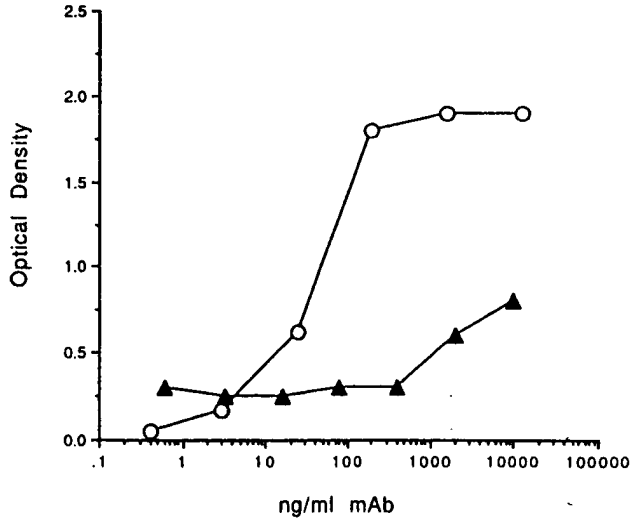


FIG. 6. Comparison of B5 mAb (triangles) and mAb A10G10 (circles) binding to soluble rhTNFα. MAbs were bound to plastic plates precoated with anti-human or anti-mouse antibody. Biotinylated TNF was then incubated with the antibodies and binding was detected by enzyme–avidin conjugates.

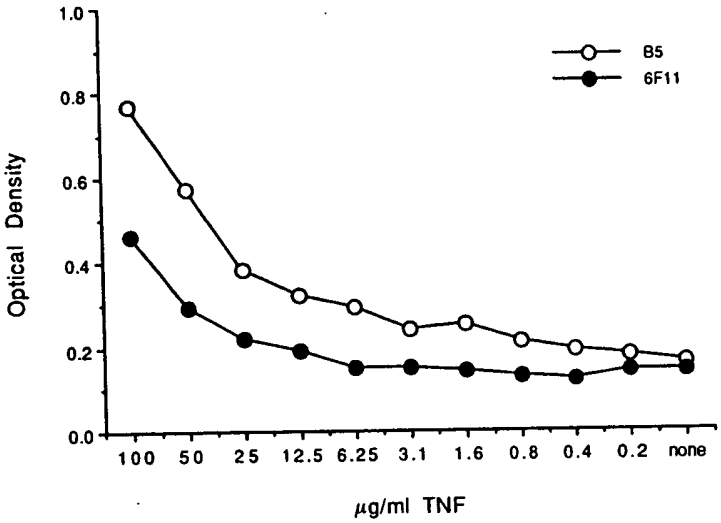


FIG. 7. Captured B5 mAb binds soluble TNFα weakly and presents it to A10G10 mAb. B5 mAb anti-TNFα or 6F11 (human anti-LPS IgM) as a control, were allowed to bind to plates precoated with anti-human IgM. Soluble TNFα was then allowed to bind to the complexed human mAbs. Mouse mAb A10G10 was added and its binding to complexed TNF was detected by enzyme-linked anti-mouse IgG antibody.

donor origin of B5 mAb is significant. The antibody is clearly different from the mouse mAbs we have generated to TNFα, all of which are neutralizing (21).

Three lines of evidence suggest that B5 mAb recognizes an epitope different from those recognized by the mouse mAbs described. First, there is no competition between the human and mouse mAbs for binding to plates coated with TNF. Second, TNF bound by the human mAb can be recognized by the mouse mAbs and vice versa. Finally, B5 mAb does not neutralize rhTNFα whereas the mouse mAbs do. One might argue that TNFα is a trimer and, as such, TNFα bound to neutralizing mouse mAbs attached to plates can still present an identical epitope to be recognized by mAb B5. The lack of competition between the mouse mAbs and mAb B5 for plate-bound TNFα is a strong argument against this possibility. The competition data in combination with the lack of neutralizing activity of B5 mAb support the interpretation of distinct epitope recognition by the mouse and human mAbs. The biological effects of TNFα, especially the ability of the cytokine to promote Ig secretion by human B cells, may preclude the generation of a high-affinity neutralizing human anti-TNFα autoantibody by the techniques used and may also explain the different epitope specificities of B5 mAb and the three neutralizing mouse mAbs.

The base of the bell-shaped trimeric TNFα molecule, which contains the amino terminus apposed to the carboxy terminus, is the region of the molecule which binds to TNF receptors (24, 25). Since the mouse mAbs used in this report neutralize TNFα, as well as block binding of TNFα to its receptors (S. Faris and C. Galloway, personal communication), it is likely that an epitope in the base of the trimer is recognized by these antibodies. The B5 mAb, and the other human IgM mAbs, can recognize trimeric and monomeric TNFα, as judged by binding in ELISA and Western blotting experiments. From the data presented in this report, one might speculate that the B5 mAb

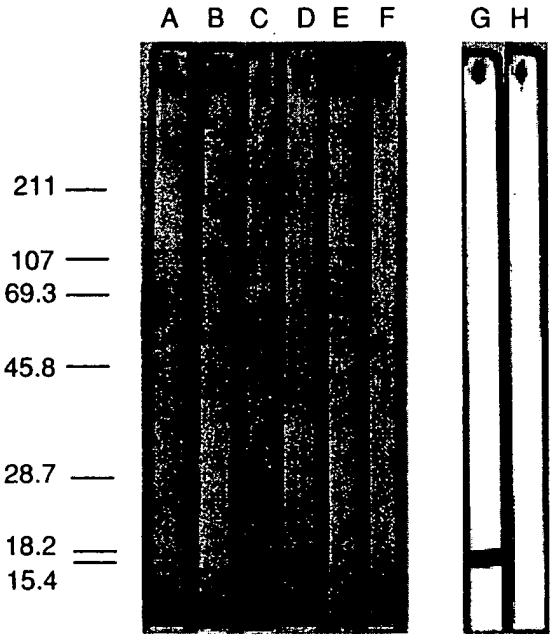


FIG. 8. Western blotting of rmTNFα (lanes A–F) and rhTNFα (lanes G and H). TNFα was electrophoresed under reducing conditions and transferred to nitrocellulose. Mouse TNFα was blotted with the following monoclonal antibodies: 7T1 (A), B5 (B), 1A6B5F (C), 6F11 (D), H5 (E), and no primary antibody (F). Lanes A–F were developed and then exposed to biotinylated anti-human IgM followed by avidin-coupled horse radish peroxidase. Human TNFα was used in lanes G and H. The mAb 6F11 was used to blot lane H and B5 was used for lane G. These were then exposed to biotinylated anti-human IgM followed by avidin-coupled horse radish peroxidase. Molecular weight standards, ranging in molecular weight from 211 to 15.4 kDa, were run in parallel and their positions are indicated.

sees a region of the TNFα molecule closer to the "top" of the trimer. The B5 mAb showed specificity for rhTNFα and some cross-reactivity with recombinant mouse TNFα. This is consistent with the observed 79% amino acid identity of these two species of TNFα (26). It is interesting, however, that several of the human mAbs analyzed bind both mouse and human TNFα. This contrasts with the observation that mouse anti-human TNFα mAbs rarely bind mouse TNFα. In fact, we are aware of only one report of a single murine anti-human TNFα mAb which also binds murine TNFα (27).

The weak binding of B5 mAb to soluble TNFα is consistent with a low binding constant of the mAb for the ligand. Nevertheless, the valency of this IgM mAb can outweigh this shortcoming so that B5 can bind to solid-phase TNFα as well as, or better than, the high-affinity neutralizing mouse anti-TNFα mAbs tested. Apparently, multipoint binding allows the mAb B5 to adhere strongly to TNFα when a sufficient antigen density is available. The ability of B5 to bind to the integral membrane protein form of TNFα, somewhat analogous to solid-phase TNFα, and the biological consequences of such binding are the subjects of a separate report.

Although B5 appears to bind with low affinity, we show that it binds specifically to TNFα and fails to bind to any of the other antigens tested. This contrasts with the observed binding of two other control polyreactive mAbs. Hence, B5 appears to be

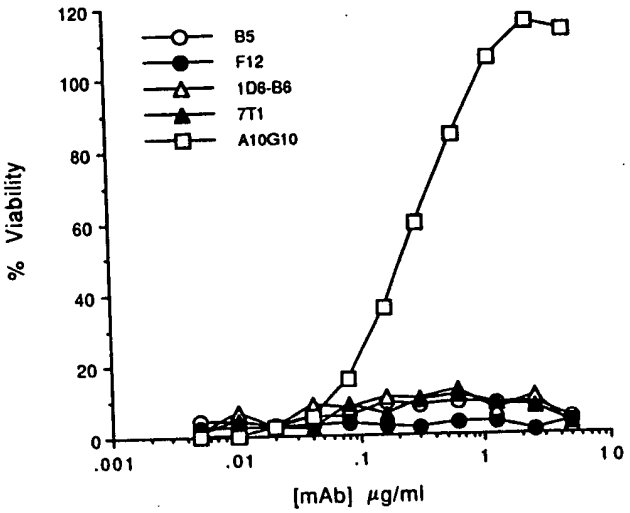


FIG. 9. Neutralization of rhTNFα by A10G10 mouse mAb and lack of neutralization by human mAbs. WEHI 164 cells were incubated with a cytotoxic dose of rhTNFα in the presence of titrated concentrations of mAb. Viability was assessed approximately 20 hr later, as described under Materials and Methods.

monospecific for an epitope shared by mouse and human TNFα. These properties classify B5 as an autoantibody and distinguish it from other mAbs so far described.

ACKNOWLEDGMENTS

We thank Ms. J. Kieran and Ms. K. Nguyen for technical assistance, Ms. S. Faris for help in establishing the TNF cytotoxicity assay, Ms. T. Wong and Dr. G. Meng for establishing the conditions for the soluble biotinylated TNF binding assay and providing biotinylated TNF, and Drs. P. Scuderi, J. Kates, J. Boyle, S. Chan, M.-S. Cho, C. Galloway, as well as Ms. D. Penza, S. Faris, J. Kieran, and Mr. R. Roby, for helpful discussions.

REFERENCES

1. Beutler, B., and Cerami, A., *Adv. Immunol.* **42,** 213, 1988.
2. Beutler, B., and Cerami, A., *Annu. Rev. Immunol.* **7,** 625, 1989.
3. Rosenblum, M. G., and Donato, N. J., *Crit. Rev. Immunol.* **9,** 21, 1989.
4. Fiers, W., *FEBS Lett.* **285,** 199, 1991.
5. Jäätelä, M., *Lab. Invest.* **64,** 724, 1991.
6. Lloyd, A. R., and Oppenheim, J. J., *Immunol. Today* **13,** 169, 1992.
7. Sung, S.-S. J., Jung, L. K. L., Walters, J. A., Chen, W., Wang, C. Y., and Fu, S. M., *J. Exp. Med.* **168,** 1539, 1988.
8. Becker, J. C., Kolanus, W., Lonnemann, C., and Schmidt, R. E., *Scand. J. Immunol.* **32,** 153, 1990.
9. Sung, S.-S. J., Bjorndahl, J. M., Wang, C. Y., Kato, H. T., and Fu, S. M., *J. Exp. Med.* **167,** 937, 1988.
10. Ranges, G. E., Zlotnik, A., Espevik, T., Dinarello, C. A., Cerami, A., and Palladino, M. A., Jr., *J. Exp. Med.* **167,** 1472, 1988.
11. Kehrl, J. H., Miller, A., and Fauci, A. S., *J. Exp. Med.* **166,** 786, 1987.
12. Jelinek, D. F., and Lipsky, P. E., *J. Immunol.* **139,** 2970, 1987.
13. Carswell, E. A., Old, L. J., Kassel, R. L., Green, S., Fiore, D., and Williamson, B., *Proc. Natl. Acad. Sci. USA* **72,** 3666, 1975.
14. Fomsgaard, A., Svensen, M., and Bendtzen, K., *Scand. J. Immunol.* **30,** 219, 1989.

15. Bendtzen, K., Svensen, M., Fomsgaard, A., Poulsen, L. K., and Hansen, M. B., *Prog. Leukocyte Biol.* **10B,** 447, 1990.
16. Leusch, H.-G., Sitzler, G., and Markos-Pusztali, S., *J. Immunol. Methods* **139,** 145, 1991.
17. Hartman, A. B., Mallett, C. P., Srinivasappa, J., Prabhakar, B. S., Notkins, A. L., and Smith-Gill, S. J., *Mol. Immunol.* **26,** 359, 1989.
18. Casali, P., and Notkins, A. L., *Immunol. Today* **10,** 364, 1989.
19. Mae, N., Liberato, D. J., Chizzonite, R., and Satoh, H., *In* "Lymphokine Cytokine and Research" Vol. 10(1), pp. 61–68. Mary Ann Liebert Publishers, New York, 1991.
20. Suzuki, H., Kamimura, J., Ayabe, T., and Kashiwagi, H., *J. Immunol.* **145,** 2140, 1990.
21. Galloway, C. J., Madanat, M. S., and Mitra, G., *J. Immunol. Methods* **140,** 37, 1991.
22. Kasaian, M. T., Ikematsu, H., and Casali, P., *J. Immunol.* **148,** 2690, 1992.
23. Meng, Y. G., and Trawinski, J., *J. Immunol.* **141,** 2684, 1988.
24. Eck, M. J., and Sprang, S. R., *J. Biol. Chem.* **264,** 17595, 1989.
25. Corti, A., Fassina, G., Marcucci, F., and Cassani, G., *Mol. Immunol.* **29,** 471, 1992.
26. Pennica, D., Hayflick, J. S., Bringman, T. S., Palladino, M. A., and Goeddel, D. V., *Proc. Natl. Acad. Sci. USA* **82,** 6060, 1985.
27. McLaughlin, P. J., Elwood, N. J., Russell, S. M., Andrew, S. M., and McKenzie, I. F. C., *Anticancer Res.* **12,** 1243, 1992.

C4

CELLULAR IMMUNOLOGY 152, 569–581 (1993)

II. The B5 Monoclonal Human Autoantibody Binds to Cell Surface TNFα on Human Lymphoid Cells and Cell Lines and Appears to Recognize a Novel Epitope

PETRA BOYLE, KENNETH J. LEMBACH, AND GAYLE D. WETZEL

Biological Products, Pharmaceutical Division, Cellular and Molecular Biology Research, Miles Inc., P.O. Box 1986, Berkeley, California 94701

Received May 28, 1993; accepted August 6, 1993

A human IgM monoclonal antibody (B5) recognizing human TNFα was established from peripheral blood lymphocytes by transformation with Epstein–Barr virus and subsequent cell fusion. The B5 monoclonal antibody (mAb) binds to cell surface TNFα (csTNFα) on human T cells, B cells, and monocytes. In addition, this autoantibody binds to csTNFα on a variety of lymphoid and monocyte lineage cell lines of human origin, as well as astrocytomas, a breast carcinoma, and a melanoma. Interestingly, the B5 mAb also binds to chimpanzee lymphocytes and to mouse T lymphoma cell line csTNFα. Many neutralizing mouse anti-TNFα mAbs do not exhibit comparable binding to csTNFα. This is consistent with the previous demonstration that B5 recognizes an epitope on TNFα distinct from those recognized by three neutralizing mouse anti-TNFα mAbs. B5 binding to csTNFα is specific since it can be inhibited by TNFα. No inhibition of B5 binding was seen by a neutralizing mouse anti-TNFα mAb. The B5 autoantibody appears to recognize the transmembrane form of TNFα and most likely also recognizes TNFα associated with its receptor. The unique specificity of this B5 autoantibody provides some additional insight into the complex physiology of cell surface-associated TNFα.

INTRODUCTION

A variety of different cell types produce and secrete TNFα/cachectin including monocytes (for reviews see 2–6), T lymphocytes (7), NK cells (8), B lymphocytes (9, 10), transformed B cell lines (9, 11), primary chronic malignant B cell isolates (12), and nonhemopoietic cells (13, 14). TNFα can also be expressed on cell surfaces (15–22). Cell-associated TNFα apparently exists in two forms. One is an integral type 2 transmembrane protein with a 26-kDa molecular weight, as shown for monocytes (16, 17) and T cells (21). The other form is the secreted 17-kDa product which is bound to its receptor (15, 23). The existence of cell surface TNFα is supported by several findings. Plasma membrane preparations of activated monocytes or paraformaldehyde-fixed intact cells, for example, demonstrate cytotoxic activity against TNFα-sensitive targets. This cytotoxicity can be blocked with antibody or neutralizing antisera to TNFα (15, 16, 19, 23). Furthermore, these cell membranes express a 26-kDa form, a 17-kDa form, or both forms of TNFα when subjected to Western blotting (15, 18, 21, 23). Activated T cells express an epitope recognized by a neutralizing anti-TNFα monoclonal antibody (mAb) (21).

0008-8749/93 $5.00

We recently described a human monoclonal anti-human TNFα antibody called B5 mAb (1). This antibody does not neutralize TNFα nor does it compete with neutralizing mouse antibodies for binding to TNFα. It thus appears to recognize a nonneutralizing epitope of TNFα. In this report, we show that this human autoantibody binds to cell surface TNFα (csTNFα) on human monocytes, B and T lymphocytes, and a variety of cell lines.

MATERIALS AND METHODS

Reagents

The mouse and human mAbs used in these experiments have been described previously (1). Briefly, B5 is a human IgM monoclonal antibody generated in our laboratory which recognizes both human and mouse TNFα. It does not neutralize TNFα cytotoxicity nor does it bind to epitopes recognized by several neutralizing mouse antibodies. A10G10 and A6 are mouse IgG_1 neutralizing anti-human TNFα mAbs and have been described previously (1). Phorbol myristic acetate (PMA), mouse IgG_1, staphylococcal enterotoxin B, and phytohemagglutinin (PHA) were purchased from Sigma (St. Louis, MO). *Escherichia coli* LPS was a gift from S. Faris (Miles Inc, Berkeley, CA). Phycoerythrin (PE) conjugated anti-CD3 and anti-CD19 antibodies were purchased from Dakopatts (Carpenteria, CA). PE-conjugated anti-CD14 (LeuM3) was purchased from Becton–Dickinson (San Jose, CA). Fluorescein (FL) conjugated $F(ab')_2$ anti-human IgM, FL–$F(ab')_2$ anti-human IgG, and FL-$F(ab')_2$ anti-mouse IgG antibodies were purchased from Cappel (Organon Teknika Corp., Durham, NC). Fetal bovine serum (FBS) was purchased from Hyclone (Salt Lake City, UT). The cell lines mentioned in Table 1 were all purchased from the ATCC (Bethesda, MD), except for the 8B9-Epstein–Barr virus (EBV)-transformed human B cell line which was obtained from Genetic Systems, Inc. The *Staphglococcus aureus* Cowan strain was purchased from Calbiochem (San Diego, CA). The anti-human IgD–dextran conjugate was a kind gift of Dr. F. Finkelman (USUHS, Bethesda, MD). Estapor uniform latex particles dyed with propidium iodide, 0.1 μm diameter, were purchased from Bangs Lab, Inc. (Carmel, IN). These were coupled with antibodies by incubating them with antibody at 1 mg/ml in bicarbonate buffer, pH 9.6, for 16 hr at 4°C. These were washed with PBS/BSA, blocked with PBS/1.5% BSA, and stored in PBS with azide and 1% FBS.

Fluorescence Analyses

One million cells were incubated with optimal concentrations of primary antibody, usually 10 μg/ml unless otherwise indicated, at 4°C for 1 hr in PBS containing 1% FBS and 0.02% sodium azide. Optimal concentrations of fluorescent secondary antibodies were added, after two cell washes, for a similar time in similar buffer. After washing, cells were fixed with 1% paraformaldehyde solution. Fluorescent latex particles conjugated with antibodies were also titrated and used at optimal concentrations. Two color analyses were performed using PE- or FL-conjugated anti-CD3, CD19, or CD14 antibodies and FL-conjugated secondary antibodies or antibodies coupled to propidium iodide-labeled particles, respectively. Cell fluorescence was analyzed on a FACSCAN (Becton–Dickinson).

RESULTS

The B5 mAb Anti-rhuTNFα Binds to the Surface of Several Different Cell Lines

Since the B5 mAb binds specifically to rhuTNFα, it was used to examine cell surface expression of TNFα by various cell lines. Fig. 1 shows the results of a typical experiment. An EBV-transformed human B lymphoblastoid cell line, 8B9, and a human monocyte cell line, THP-1, were incubated with either B5 anti-TNFα mAb or, as a control, the 6F11 anti-pseudomonas LPS mAb and then fluorescent anti-human IgM $F(ab')_2$ fragments. The 8B9 cells showed significant binding of the B5 mAb whereas no significant binding to the cell surface was seen with the control 6F11 mAb. B5 binding was also

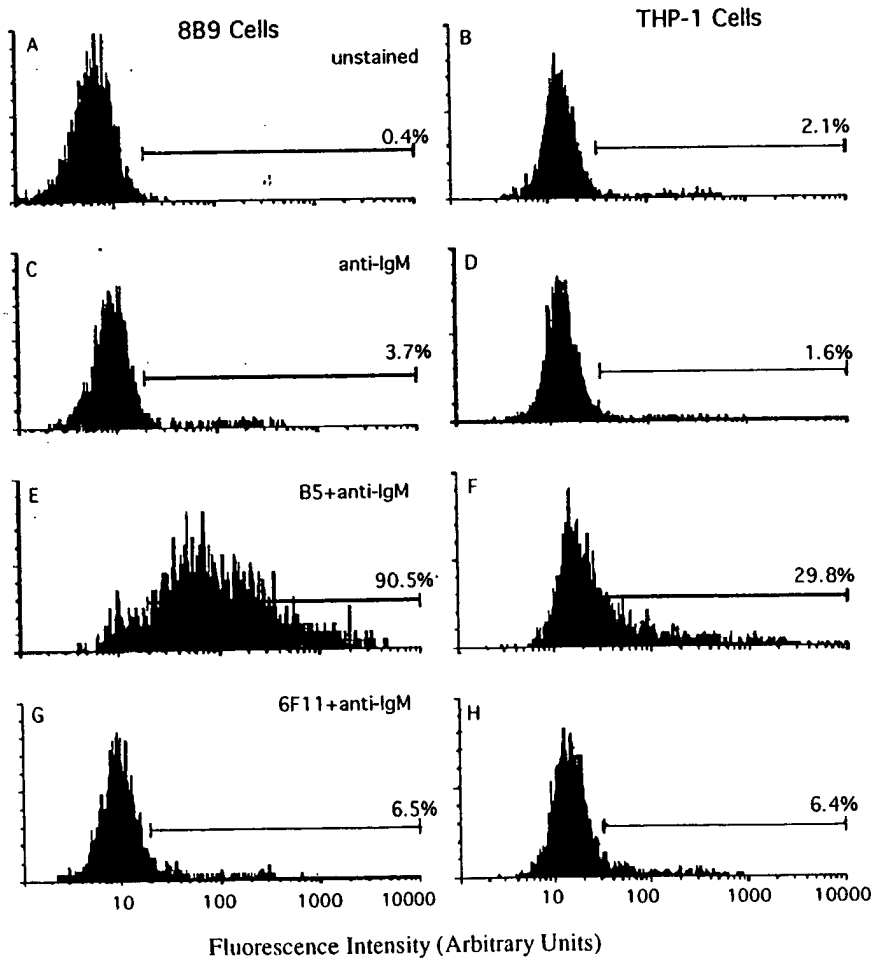


FIG. 1. Fluorescence histograms of two cell lines incubated with human IgM anti-TNFα mAbs. 8B9 cells (A, C, E, G) and THP-1 cells (B, D, F, H) were incubated with no antibodies (A, B), with FL-$F(ab')_2$ anti-human IgM (C, D), B5 IgM anti-TNFα + FL-anti-IgM (E, F), and 6F11 anti-LPS + FL-anti-IgM (G, H). Fluorescence intensity, in arbitrary units, is plotted against the cells per channel on the ordinate. For each sample 5000 cells were accumulated. Percentages of cells falling within the indicated markers, scored as fluorescence positive, are given.

observed with THP-1 cells. However, fewer cells in this population exhibited binding and the intensity of the fluorescent signal was somewhat lower than that seen for the 8B9 cells. Nearly one-third of the cells in the THP-1 population expressed csTNFα, as detected with the B5 mAb. It is unclear whether this level of B5 binding reflects some regulation of csTNFα expression or whether it is due to clonal variation within the cell line.

Detection of csTNFα was examined more closely with the THP-1 monocyte and U937 histiocyte cell lines. These cells were incubated for 3 hr with no stimulus, LPS, or LPS + PMA and were then incubated with titrated amounts of B5 antibody. The results are shown in Fig. 2. In all cases, B5 binding to cells was dose dependent. Interestingly, more binding was observed for both cell lines when they were preincubated with LPS or LPS + PMA. This was especially apparent for the U937 cell line. Upon stimulation, B5 binding to the cells was readily apparent, even using low concentrations of antibody. These results are consistent with the known ability of LPS and PMA to induce TNF secretion by monocyte cell lines.

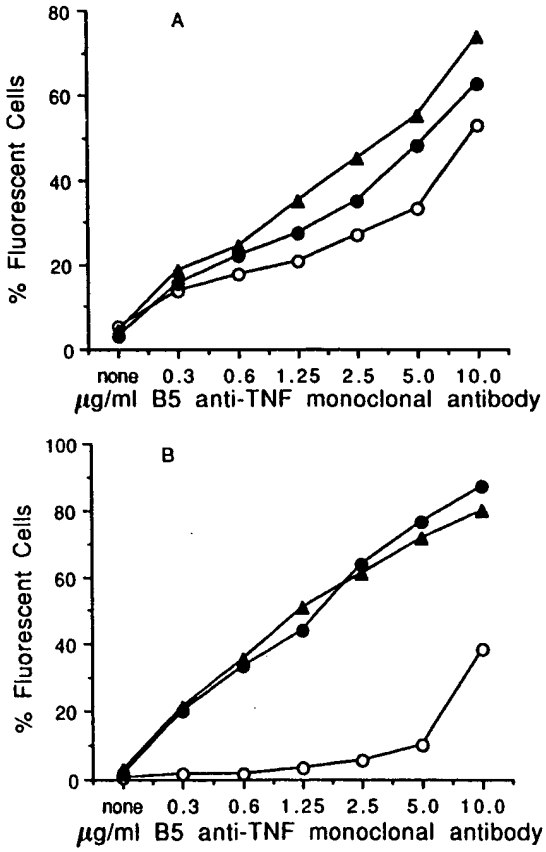


FIG. 2. Detection of cell surface expression of TNFα on THP-1 and U937 cells with the B5 anti-TNFα mAb and increase in expression with LPS and PMA. THP-1 (A) and U937 (B) cells were incubated 3 hr with medium (open circles), LPS (filled circles), or LPS + PMA (filled triangles). Fluorescence was determined with a FACSCAN.

Table 1 shows experiments in which csTNFα expression was surveyed using a variety of human and mouse cell lines. Surprisingly, B5 bound to cells derived from both species. Many of the cell lineages tested have been reported to secrete TNFα. Of the 17 lines tested, only the breast carcinoma U118MG line showed no binding by B5. The other cell lines exhibited a range in the percentages of cells within each population which expressed csTNFα from a low of around 8% for the A375 melanoma to over 90% for EBV-transformed B cells. The class-matched 6F11 anti-LPS mAb failed to bind to any of these cell lines. This observation and the negative results with U118MG indicate that the B5 binding seen was specific and not the result of a general affinity for all cells.

Lack of a Neutralizing Mouse Anti-TNFα mAb Binding to csTNFα

ELISA experiments have shown the TNF specificity of the B5 mAb and demonstrated binding to an epitope on TNFα different from that recognized by the neutralizing mouse mAb A10G10. We next examined whether the epitope recognized by A10G10 was expressed on the surface of cells to which B5 binds. Table 2 presents data from five experiments addressing this issue using the U937 and THP-1 cell lines. In all five experiments the B5 mAb bound to each cell line. On the other hand, A10G10

TABLE 1

Binding of the TNFα-Specific B5 Human IgM mAb to Various Cell Lines

				% Fluorescent cells		
Expt	Cell line	Species	Cell type	Primary antibody: None	B5	6F11
1	8B9-EBV	Human	B lymphoblast	1.1	86.9	4.7
	1A2-EBV	Human	B lymphoblast	2.3	64.7	2.7
	hpbl-EBV	Human	B lymphoblast	2.0	96.2	2.3
	cpbl-EBV	Chimpanzee	B lymphoblast	6.6	76.1	6.2
	Tonsil-EBV	Human	B lymphoblast	4.6	91.2	4.9
	Jurkat	Human	T lymphoma	0.7	17.9	1.2
	LBRM33	Mouse	T lymphoma	3.1	72.7	3.8
	$P3X63_{Ag8.653}$	Mouse	Myeloma	6.8	78.4	5.6
	DU4475	Human	Breast carcinoma	10.2	52.4	9.8
	SW1088	Human	Astrocytoma	11.2	15.3	10.9
	U118MG	Human	Glioblastoma	6.2	7.3	6.2
	U373	Human	Glioblastoma/ Astrocytoma	4.9	69.6	3.5
2	U937	Human	Histiocytic lymphoma	0.9	63.1	1.5
	THP-1	Human	Monocyte	1.7	25.2	2.0
	1A2-EBV	Human	B lymphoblast	2.2	98.4	2.9
	8B9-EBV	Human	B lymphoblast	4.7	98.8	5.4
	A375	Human	Melanoma	1.7	8.5	2.5
3	HL-60	Human	Promyelocyte	8.0	41.2	6.6
	J774A.1	Mouse	Monocyte/mϕ	5.4	82.0	4.8

Note. Cells were incubated with the indicated primary antibodies and fluorescein-labeled anti-human IgM (μ-specific) secondary antibody. The percentages of cells exhibiting fluorescence as determined on a FACSCAN are shown.

TABLE 2

Binding of Human B5 and Lack of Binding of Mouse A10G10 Anti-TNFα mAbs to Cell Surface TNF on Unstimulated Monocyte and Histiocyte Cell Lines

Expt	Cell line	% Fluorescent cells					
		Primary antibody: None	A10G10	$mIgG_1$	None	B5	6F11
1	U937	0.3	0.4	nd	2.9	15.1	2.7
2	THP-1	0.9	2.6	nd	2.3	24.8	2.8
	U937	0.8	2.4	nd	2.7	99.1	2.8
3	THP-1	3.1	2.7	nd	2.7	34.7	3.4
	U937	1.6	1.9	nd	1.6	35.7	1.8
4	THP-1	2.4	3.1	1.6	1.7	17.7	3.4
	U937	2.8	2.9	2.8	2.2	20.2	2.7
5	THP-1	4.7	6.0*	6.7	4.6	56.3	nd
	U937	1.5	9.9*	2.3	1.2	61.4	nd

Note. Cells were incubated with the indicated primary antibodies and fluorescein-labeled anti-mouse IgG (γ-specific) or anti-human IgM (μ-specific) secondary antibodies. Asterisks (*) indicate that $F(ab')_2$ fragments of A10G10 mAb were used. The percentages of cells exhibiting fluorescence as determined on a FACSCAN are shown. Not determined is signified by nd.

did not bind to a significant degree in four of the experiments. In one experiment, a small amount of binding by A10G10 to the U937 cells was observed. Taken together, these data suggest that epitopes of TNFα recognized by B5 are constitutively expressed on the surface of these cell lines, but the epitopes recognized by A10G10 are only rarely expressed.

LPS Induction of Cell Surface TNFα Expression

We next incubated THP-1 and U937 cells with LPS to examine whether or not expression of different csTNFα epitopes can be increased. Table 3 shows the results of three experiments. In all three, LPS increased the amount of B5 binding to THP-1 cells. This was true also for U937 cells in two of the three experiments. Interestingly, LPS stimulation resulted in cell surface binding of the A10G10 mAb for both the THP-1 and the U937 lines. The range of the increased number of cells binding A10G10 after LPS stimulation was between about 5 and 14%, depending on the experiment. The corresponding increases observed with B5 binding ranged from about 8 to about 11%. These data suggest that csTNFα can be increased by incubation with LPS and that this increase correlates with the acquisition of TNFα epitopes recognized by neutralizing antibodies.

Specificity of B5 mAb Binding to csTNFα

Table 4 presents data which confirm the specificity of B5 mAb binding to the THP-1 cells. Preincubation of the IgM mAb with TNFα was found to inhibit subsequent cell surface binding in a dose-dependent manner. Except for a small effect at the highest concentration, lymphotoxin (TNFβ) did not inhibit B5 binding. The lack of

TABLE 3

Analysis of Cell Surface Expression of TNFα after Induction with Lipopolysaccharide

Expt	Cell line	LPS	% Fluorescent cells					
			Primary antibody: None	A10G10	$mIgG_1$	None	B5	6F11
1	THP-1	−	3.1	2.7	nd	2.7	34.7	3.4
		+	6.9	16.5	nd	3.7	43.8	3.6
	U937	−	1.6	1.9	nd	1.6	35.7	1.8
		+	3.9	12.7	nd	1.9	43.5	2.6
2	THP-1	−	2.4	3.1	1.6	1.7	17.7	3.4
		+	3.1	8.3	2.2	3.0	29.6	3.2
	U937	−	2.8	2.9	2.8	2.2	20.2	2.7
		+	3.6	11.8	2.4	2.4	28.4	3.2
3	THP-1	−	4.7	6.0*	6.7	4.6	56.3	nd
		+	8.1	4.9*	5.4	5.0	65.9	nd
	U937	−	1.5	9.9*	2.3	1.2	61.4	nd
		+	1.0	13.1*	3.7	0.7	49.3	nd

Note. Stimulation was performed by 3- or 4-hr incubation with 10 μg/ml LPS. Cells were incubated with the indicated primary antibodies and fluorescein labeled anti-mouse IgG (γ-specific) or anti-human IgM (μ-specific) secondary antibodies. Asterisks (*) indicate that $F(ab')_2$ fragments of A10G10 mAb were used. The percentages of cells exhibiting fluorescence as determined on a FACSCAN are shown. Not determined is signified by nd.

complete inhibition with high doses of TNFα is consistent with the previously documented low affinity of the B5 mAb for soluble TNFα. Interestingly, A10G10 had no effect on the binding of B5 mAb to THP-1 cells. These data suggest that the neutralizing mAb A10G10 does not compete for the same epitope on TNFα to which B5 mAb binds. Furthermore, inhibition of B5 binding by TNFα indicates the specificity of the binding.

B5 Binds to csTNFα on Fresh Human Spleen Cells

The previous sections establish B5 binding to csTNFα on several different cell lines. To determine whether B5 binds to untransformed cells, experiments were performed with human splenocytes.

To analyze B cell expression of csTNFα by B5, we used unconjugated B5 IgM since direct fluoresceination or biotinylation of this antibody was very inefficient or interfered with its TNFα binding ability. Fluorescent $F(ab')_2$ fragments of anti-human IgM antibody were used to detect B5 binding. Since many normal B cells already express sIgM as an antigen receptor, it was not always possible to detect csTNFα as an increase in the percentage of sIgM+ cells. We could, however, detect csTNFα by measuring the increase in staining intensity with the fluorescent anti-IgM when cells were incubated with the B5 mAb compared to incubation with either control 6F11 IgM mAb or no antibody at all. Figure 3 demonstrates the shift in fluorescence intensity seen when the B5 mAb was incubated with B cells. Figure 3C shows the fluorescence histogram of cells incubated with anti-IgM antibody alone. Figure 3B shows a histogram of the same cells when first reacted with B5 mAb and subsequently incubated with

TABLE 4

TNFα Inhibition of B5 Anti-TNFα mAb Binding to THP-1 Cell Surface

	% Fluorescent cells				
Inhibitor	μg/ml Inhibitor: 0.0	0.03	0.30	3.0	30.0
TNFα	44.1	43.2	35.9	22.2	15.8
LT	44.1	39.8	40.0	40.0	29.7
A10G10	44.1	39.6	40.9	44.4	41.9

Note. B5 mAb at 10 μg/ml was incubated with the indicated concentrations of inhibitors prior to exposure to LPS-stimulated THP-1 cells. Binding of B5 was detected with fluorescein-conjugated $F(ab')_2$ anti-human IgM antibody. LT is recombinant human lymphotoxin, and A10G10 is the neutralizing mouse IgG_1 anti-TNFα mAb. Analyses were performed with a FACSCAN.

fluorescent anti-IgM antibody. The most useful parameter to measure this shift is the median channel of fluorescence intensity, or simply median channel.

Table 5 presents the data from three experiments using splenic biopsy material. In Experiments 1 and 2, the expression of csTNFα on monocytes, T cells, and B cells was examined by two-color immunofluorescence analysis using PE-conjugated anti-CD14, anti-CD3, and anti-CD19, respectively, in conjunction with FL-conjugated anti-human IgM. In Experiment 3, antibodies coupled to propidium iodide labelled beads were used in conjunction with FL-conjugated anti-CD3, anti-CD19, and anti-CD14 mAbs. This removed any ambiguity arising from the use of FL–anti-IgM as a secondary reagent.

Although monocytes constituted less than 5% of the total splenocyte populations in all experiments, the anti-TNFα B5 mAb bound to a significant fraction. On the other hand, monocytes did not bind the control 6F11 human IgM mAb. These results suggest that some splenic monocytes express csTNFα.

CD3+ T cells showed variable expression of csTNFα. While the percentages of csTNFα-positive T cells varied in these experiments, the fluorescence intensity observed with the B5 mAb was much weaker than that seen for B cells and monocytes. The median fluorescence intensity for T cell csTNFα was less than twice that seen for the background controls. These results suggest that a variable proportion of splenic T cells express small amounts of csTNFα.

In contrast, analysis of B cells indicated high-level expression of csTNFα. Furthermore, the neutralizing mouse anti-human TNFα mAb A6 bound to some B cells in Experiment 3. No binding was seen with the 6F11 control antibody, indicating the specificity of the B5 interaction with B cells. These data suggest that the B5 anti-TNFα mAb can react with splenic lymphocytes of the B and T lineages as well as being able to recognize and bind to splenic monocytes.

B5 Binding to csTNFα on Cultured Human Spleen Cells

Spleen cells from two of the individuals examined in Table 5 were cultivated *in vitro* with various stimuli and were then analyzed for B5 mAb binding. The results from both experiments were similar and the data from one, for clarity, are shown in Table 6. Cultivation of the cells resulted in loss of monocytes, therefore, data for CD14+ cells are not presented.

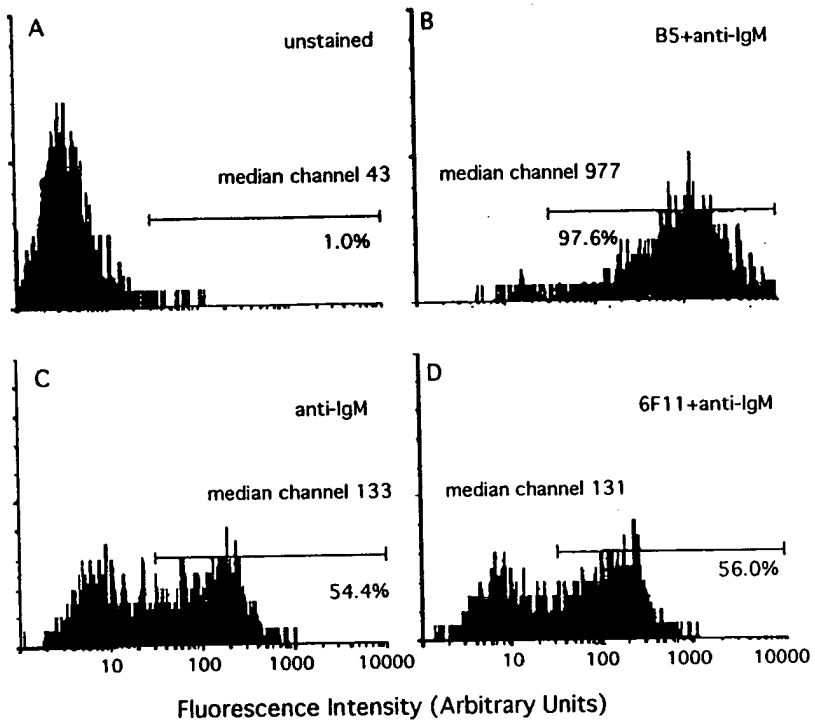


FIG. 3. Binding of B5 anti-TNFα IgM mAb to CD19-positive splenic B cells. Splenocytes were incubated with phycoerythrin-conjugated anti-CD19 and only positive cells were further analyzed for fluorescein-conjugated antibody binding. A shows C19+ splenocytes not incubated with FL-anti-IgM. B shows incubation of these cells with B5 + FL-anti-IgM, C shows incubation with FL-anti-hIgM alone, and D shows incubation with control 6F11 anti-LPS IgM + FL-anti-IgM. The percentages of cells within the indicated markers are given, indicating the percentage of cells reacting positively with the fluorescein-conjugated antibody. The median channel numbers for the positive populations are also given. These numbers reflect the fluorescence intensity, measured in arbitrary units, for the fluorescence-positive populations.

Cultures containing the T cell activator PHA showed B5 binding to both T cells and B cells. Similarly, cultures containing the B cell activators anti-IgD-dextran + IL-2 showed B5 binding to both B cells and T cells. In these cultures, IL-2 may stimulate some preactivated T cells and account for the increased B5 binding to CD3+ cells compared to that seen for resting T cells in Table 5. Binding by the control mAb 6F11 was not observed, suggesting that neither the B5 nor the 6F11 mAbs bound to IgM receptors expressed on activated T cells. These putative IgM receptors (24) are presumably already occupied by IgM secreted into the medium. These data suggest that both activated human B cells and T cells express csTNFα epitopes recognized by the B5 mAb.

B5 Binding to csTNFα on Peripheral Blood Lymphocytes

Human peripheral blood lymphocytes from two individuals were examined and Table 7 shows the results. Significant fractions of T cells, B cells, and monocytes were bound by the B5 mAb in both experiments. The B cells exhibited more pronounced

TABLE 5

Analysis of Cell Surface TNFα Expression on Fresh Human Splenocytes

Expt	Cells analyzed	% Fluorescent cells (median fluorescence intensity channel) First Ab: None / Second Ab: Anti-IgM	B5 Anti-IgM	6F11 Anti-IgM
1	Spleen 1			
	CD3+	3.4 (41)	20.9 (24)	3.6 (53)
	CD14+	4.4 (71)	23.2 (78)	4.7 (104)
2	Spleen 2			
	CD3+	5.9 (39)	80.8 (57)	4.0 (42)
	CD19+	62.9 (122)	98.1 (906)	63.8 (121)
	CD14+	8.2 (194)	49.2 (172)	7.7 (184)

Expt		No beads	Antibody coupled to beads: 6F11	B5	None	A6	A10G10
3	Spleen 3						
	CD3+	0.3	2.8	4.6	2.2	4.5	1.9
	CD19+	1.5	5.2	40.4	2.7	7.0	2.4
	CD14+	1.1	2.3	31.3	0.7	3.3	0.7

Note. Human splenocytes received 1 day after biopsy were analyzed for expression of cell surface binding with the indicated monoclonal antibodies. T cells, B cells, and monocytes were identified with PE- (Experiments 1 and 2) or FL-conjugated (Experiment 3) anti-CD3, anti-CD19, and anti-CD14 (LeuM3) antibodies, respectively. Two-color analyses were then performed on these populations using FL-labeled $F(ab')_2$ anti-human IgM and the indicated IgM mAbs (Experiments 1 and 2) or propidium iodide-labeled latex beads (Experiment 3) coupled with the indicated mAbs. Underlined values represent those which show significant increases in the percentage of fluorescent cells or show greater than twice the fluorescence intensity of the appropriate control population. Analyses were performed with a FACSCAN.

binding. Variable proportions of monocytes bound the mouse A6 and A10G10 mAbs. The results of similar experiments with peripheral blood from two chimpanzees yielded similar results. The B5 mAb bound to significant fractions of B cells, T cells, and monocytes from these animals (data not shown). These results suggest that only some normal human and chimpanzee peripheral blood monocytes and T lymphocytes express csTNFα whereas a larger fraction of total peripheral B lymphocytes express this cell surface cytokine.

DISCUSSION

We have shown that the human B5 autoantibody binds to TNFα on a broad range of human cell lines and lymphoid cells. As there is no difference in the amino acid sequences of chimpanzee and human TNFα, B5 binding to chimpanzee cells is predictable. More unusual, however, is the binding of B5 to csTNFα on mouse cell lines since few anti-TNF mAbs bind TNF from both species. B5 appears to be an exception since it binds to secreted mouse TNF (1) and to csTNFα, as demonstrated here. Others have certainly described TNF production by human B cells (9, 10) T cells (7), monocytes (26), B cell lines (9, 11), astrocytes (25–27), as well as some TNF-resistant cell

TABLE 6

Analysis of Cell Surface TNFα Expression by Cultured Human Splenic Lymphocytes

		% Fluorescent cells					
			Antibody coupled to beads				
Activator	Cells	No beads	6F11	B5	None	A6	A10G10
PHA	CD3+	4.7	4.2	24.9	6.0	8.1	6.3
	CD19+	2.4	3.7	38.0	3.6	5.7	3.7
anti-IgD–dextran plus IL-2	CD3+	6.8	5.4	14.0	6.0	7.2	5.7
	CD19+	4.4	5.2	78.2	4.4	10.0	5.8

Note. Human splenic lymphocytes were cultured 2 days with the indicated activators. In both experiments two-color analyses were performed. Fluorescein-conjugated anti-CD3 and anti-CD19 antibodies were used with propidium iodide-labeled latex beads (0.1 μm diameter) coupled with the indicated antibodies. Analyses were performed with a FACSCAN.

lines (14). We extend these findings to include at least one metastatic breast carcinoma, DU4475, a melanoma, A375, and the U373 astrocytoma/glioblastoma. Expression of csTNFα on human splenic lymphoid cells is also demonstrated. This is somewhat surprising since activated cells were previously employed by others to establish the presence of csTNFα. It is possible many of the cells we examined were partially activated or at a stage of differentiation where they could express this cell surface molecule. The smaller percentages of T lymphocytes and monocytes from human peripheral blood expressing csTNFα is consistent with the resting phenotype of these cells. In any case, the breadth of csTNFα expression suggests it has an important role on the surface of many cells.

TABLE 7

Analysis of Human Peripheral Blood Mononuclear Cell Surface Expression of TNFα

		% Fluorescent cells					
			Antibody coupled to beads				
Expt	Cells	No beads	6F11	B5	None	A6	A10G10
1	CD3+	0.7	1.3	20.6	2.3	3.4	3.0
	CD19+	1.1	1.6	95.9	2.5	7.1	2.7
	CD14+	0.9	2.1	46.6	5.8	30.6	5.9
2	CD3+	nd*	2.1	15.4	nd	6.7	3.7
	CD19+	nd	2.0	48.7	nd	8.5	2.6
	CD14+	nd	1.7	14.7	nd	6.8	5.1

Note. Peripheral blood mononuclear cells were prepared by separation on Ficoll and incubated with or FL-derivatized anti-CD3, CD19, or CD14. Two-color analyses were performed with a FACSCAN using propidium iodide-labeled latex beads coupled with the indicated mAbs. Underlined values represent those which show significant increases in the percentage of fluorescent cells.

Others have shown that TNFα can exist as both an integral transmembrane protein and as a mature protein bound to its receptor on cell surfaces (23). Several observations suggest that the B5 mAb recognizes primarily the integral transmembrane protein. B5 binding was increased when cells were activated with LPS or PMA. Both agents, but especially PMA, down-regulate TNF receptor expression on a variety of cell types (28, 29). B5 binds to unstimulated cell lines whereas cell lines normally need to be induced to secrete TNF. Hence, unstimulated cell lines would be expected to display little receptor-bound TNF. We showed that B5 binding to cell surfaces was inhibited by preincubation with TNFα, but not with A10G10 anti-TNFα mAb. These data demonstrate the specificity of the B5 antibody and are most consistent with binding to the integral transmembrane protein.

TNFβ binds to the same receptors as TNFα and might thereby dissociate some receptor-bound TNFα on cell surfaces. Since high doses of TNFβ were found to decrease binding of B5 to THP-1 cells (Table 4), the mAb may also recognize receptor-bound TNF.

One interesting aspect of these studies is that the human mAb B5 generally binds better to csTNF than the high-affinity neutralizing murine mAb A10G10. It is clear that the two antibodies recognize nonoverlapping epitopes (1). Western blotting experiments also suggest that A10G10 does not bind to TNFα monomers and probably recognizes a conformational epitope (data not shown). In contrast, B5 binds to monomeric and denatured forms of TNFα and likely recognizes a linear epitope. Since A10G10 neutralizes TNFα cytotoxicity (30) and prevents TNFα interaction with receptor, the antibody probably binds to TNFα near the receptor binding domain. Others have shown that mAbs which bind the amino-terminal domain of TNFα block cytotoxicity (31). Therefore, it is reasonable to postulate that A10G10 binds to amino acid residues near the amino terminus of 17-kDa TNFα, which are most membrane proximal in the transmembrane form of the cytokine. The epitope for A10G10 may thus be less accessible to antibody in csTNF than that recognized by B5. The A6 mAb, however, does occasionally bind to cells. This mAb does bind to denatured TNF in Western blots (data not shown) and has an epitope specificity distinct from that of A10G10 (1, 32).

It is interesting to note that LPS stimulation increased A10G10 binding to cells. Such stimulation is known to elicit secretion of mature trimeric TNFα, which may then bind to available TNF receptors. Since the trimer is multivalent, TNFα may bind to some receptors in a conformation which maintains either or both of the remaining receptor binding domains accessible to antibody. This form of receptor-bound TNFα may be the cell surface form of the cytokine recognized by A10G10 in LPS-stimulated cells. In support of this proposal, others have shown that incubation of nonactivated paraformaldehyde-fixed human monocytes with TNFα results in cytokine binding to receptors in a biologically active (i.e., cytotoxic) form. The cytotoxic activity of such fixed cells is inhibited by neutralizing anti-TNFα antibodies (33).

One model which is consistent with the data presented assumes that transmembrane csTNF exists as a monomer, which upon cell activation is polymerized to a trimer. Trimerization of csTNFα with a corresponding conformational change is then proposed to allow a proteolytic cleavage and release of mature biologically active TNFα trimer. B5 is considered to bind to a membrane distal domain of the csTNFα monomer. This monomer most likely has a conformation different from that of the trimer and so may expose some epitopes distinct from those displayed by the trimer. This model is

consistent with the greater binding of B5 to csTNFα. The model and data predict that more cells express cell surface TNF monomers than trimers, or that more monomer is expressed on the surface per cell than the trimer.

ACKNOWLEDGMENTS

We thank Drs. P. Scuderi, J. Kates, J. Boyle, S. Chan, M.-S. Cho, C. Galloway, G. Meng, as well as Ms. D. Penza, S. Faris, T. Wong, L. Tsuchiyama, J. Kieran, and Mr. R. Roby for helpful criticism and discussion.

REFERENCES

1. Boyle, P. A., Lembach, K. J., and Wetzel, G. D., *Cell. Immunol.* **152,** 000–000, 1993.
2. Beutler, B., and Cerami, A., *Adv. Immunol.* **42,** 213, 1988.
3. Beutler, B., and Cerami, A., *Annu. Rev. Immunol.* **7,** 625, 1989.
4. Rosenblum, M. G., and Donato, N. J., *Crit. Rev. Immunol.* **9,** 21, 1989.
5. Fiers, W., *FEBS Lett.* **285,** 199, 1991.
6. Jäätelä, M., *Lab. Invest.* **64,** 724, 1991.
10. Smeland, E. B., Blomhoff, H. K., Funderud, S., Shalaby, M. R., and Espevik, T., *J. Exp. Med.* **170,** 1463, 1989.
11. Jochems, G. J., Klein, M. R., Jordens, R., Pascual-Salcedo, D., van Boxtel-Oosterhof, F., van Lier, R. A. W., and Zeijlemaker, W. P., *Hum. Antibodies Hybridomas* **2,** 57, 1991.
12. Cordingly, F. T., Hoffbrand, A. V., Heslop, H. E., Turner, M., Bianchi, A., Reitte, J. E., Vyakarnam, A., Meager, A., and Brenner, M. K., *Lancet* **April 30,** 969, 1988.
13. Lloyd, A. R., and Oppenheim, J. J., *Immunol. Today* **13,** 169, 1992.
14. Rubin, B. Y., Anderson, S. L., Sullivan, S. A., Williamson, B. D., Carswell, E. A., and Old, L. J., *J. Exp. Med.* **164,** 1350, 1986.
15. Bakouche, O., Ichinose, Y., Heicappell, R., Fidler, I. J., and Lachman, L. B., *J. Immunol.* **140,** 1142, 1988.
16. Decker, T., Lohmann-Matthes, M.-L., and Gifford, G. E., *J. Immunol.* **138,** 957, 1987.
17. Kriegler, M., Perez, C., DeFay, K., Albert, I., and Lu, S. D., *Cell* **53,** 45, 1988.
18. Perez, C., Albert, A., DeFay, K., Zachariades, N., Gooding, L., and Kriegler, M., *Cell* **63,** 2512, 1990.
19. Espevik, T., and Nidssen-Meyer, J., *Immunology* **61,** 443, 1987.
20. Liu, C. C., Detmers, P. A., Jiang, S., and Young, J. D.-E., *Proc. Natl. Acad. Sci. USA* **86,** 3286, 1989.
21. Kinkhabwala, M., Sehajpal, P., Skolnik, E., Smith, D., Vlassara, H., Sharma, V. K., Cerami, A., and Suthanthiran, M., *J. Exp. Med.* **171,** 941, 1990.
22. Chensue, S., Remick, D., Shmyr-Forsch, C., Beals, T., and Kunkel, S., *Am. J. Pathol.* **133,** 564, 1988.
23. Luettig, B., Decker, T., and Lohmann-Matthes, M.-L., *J. Immunol.* **143,** 4034, 1989.
24. Moretta, L., Ferrarini, M., Durante, M. L., and Mingari, H., *Eur. J. Immunol.* **5,** 565, 1975.
25. Lieberman, A. P., Pitha, P. M., Shin, H. S., and Shin, M. L., *Proc. Natl. Acad. Sci. USA* **86,** 6348, 1989.
26. Chung, I. Y., and Benveniste, E. N., *J. Immunol.* **144,** 2999, 1990.
27. Selmaj, K., Raine, C. S., Cannella, B., and Brosnan, C. F., *J. Clin. Invest.* **87,** 949, 1991.
28. Ding, A. H., Sanchez, E., Srimal, S., and Nathan, C. F., *J. Biol. Chem.* **264,** 3924, 1989.
29. Aggarwal, B. A., and Eessalu, T. E., *J. Biol. Chem.* **262,** 16450, 1987.
30. Galloway, C. J., Madanat, M. S., and Mitra, G., *J. Immunol. Methods* **140,** 37, 1991.
31. Socher, S. H., Riemen, M. W., Martinez, D., Friedman, A., Tai, J., Quintero, J. C., Garsky, V., and Oliff, A., *Proc. Natl. Acad. Sci. USA* **84,** 8829, 1987.
32. Bloom, J., *J. Chromatogr.* **574,** 219, 1992.
33. Nii, A., and Fidler, I. J., *Lymphokine Res.* **9,** 113, 1990.

Journal of Immunological Methods, 139 (1991) 145–147

ADONIS 0022175991001701

JIM 05932

Letter to the editors

Failure to demonstrate TNFα-specific autoantibodies in human sera by ELISA and Western blot

Hans-Georg Leusch, Gisela Sitzler and Susanne Markos-Pusztai

Institute of Medical Immunology, Medical Faculty, Pauwelsstrasse 30, 5100 Aachen, F.R.G.

(Received 24 October 1990, revised received 24 January 1991, accepted 15 February 1991)

Dear Editors,

As recently described (Bendtzen et al., 1989, 1990; Fomsgaard et al., 1989) autoantibodies to TNFα may be detected by Western blot analysis in the sera of patients with Gram-negative septicemia, rheumatoid arthritis and in some healthy individuals. The specificity of the autoantibodies may be established by absorption of the sera with recombinant TNFα prior to the assay. Using a TNFα cytotoxicity bioassay (cell line L929) Seckinger et al. (1989) detected TNFα inhibitory factors in urine and Foley et al. (1990) found similar activity in sera of patients with sarcoidosis, tuberculosis and Crohn's disease. The specificity of the TNFα inhibitor in urine samples was shown by neutralizing the inhibitory activity with increasing amounts of human recombinant TNFα. In contrast to the findings with human urine Foley et al. (1990) found that there was no correlation between the inhibitory effect of the sera in the L929 bioassay and the ability of the sera to compete with the binding of an enzyme-conjugated TNFα monoclonal antibody to immobilized TNFα in an ELISA system. Therefore, these results contradict the findings of Fomsgaard and the existence of specific autoantibodies binding directly to TNFα remains, at the very least, controversial.

Correspondence to: H.G. Leusch, Institute of Medical Immunology, Medical Faculty, Pauwelsstrasse 30, 5100 Aachen, F.R.G.

Proceeding on the assumption that autoantibodies to cytokines such as TNFα may have a regulatory effect in the inflammatory cascade we have sought evidence for specific antibodies to TNFα in several immunoglobulin preparations. In addition, we analysed different human sera collected from patients with septicemia ($n = 11$), with chronic autoimmune disorders ($n = 10$) and from renal tumor patients ($n = 4$) undergoing treatment with rIFNα and rTNFα. ELISA and Western blot assays were used for these investigations. To test for ELISA reactivity microtiter plates (Nunc, immunoplates) were coated with 100 μl of monoclonal mouse anti-human rTNFα (clone 199, 1 : 20, Boehringer, Mannheim, F.R.G.) overnight at room temperature in 50 mM carbonate buffer pH 9.6 followed by blocking the wells with 1% BSA in the same buffer for 1 h at room temperature. Binding of the human TNFα (10 ng/ml) (Knoll, Mannheim, F.R.G.) in incubation buffer (50 mM Hepes, 150 mM NaCl, 200 mM sodium potassium tartrate, 1% BSA, 0.5% Synperonic F68, 0.75% PEG 40,000, 0.01% phenol, pH 7.0) was performed at 4°C for 18 h. After washing the plates three times with PBS-Tween (0.1%) 100 μl of patients' sera, or the corresponding protein G isolated IgG fractions, or 100 μl of immunoglobulin preparation (final concentration of 4 mg IgG/ml) were added and incubated overnight at 4°C. Afterwards, wells were washed again, filled with 100 μl of a rabbit anti-human IgG-alkaline phosphatase conjugate (Sigma, Deisenhofen, F.R.G.) (1/2000 in incubation buffer) and incubated at room temperature

for 2 h. Color was developed with 100 µl *o*-nitrophenyl-phosphate in substrate buffer (Behring, Marburg, F.R.G.). The reaction was stopped with 100 µl 1 M NaOH and absorbance measured at 405/492 nm. The results were evaluated by comparing the absorbance values of the patients' sera with the negative control lacking rTNFα. Binding of rTNFα to the immobilized mouse anti-TNFα was tested using a rabbit polyclonal anti-rTNFα (Genezyme, Boston, U.S.A.). Binding was detected using a goat anti-rabbit alkaline-phosphatase conjugated antibody (Sigma, Deisenhofen, F.R.G.) at a dilution of 1/2000 in incubation buffer. In addition, we measured the TNFα concentration with a monoclonal TNFα-ELISA procedure (anti-TNFα MAK clone 199 and anti-TNFα clone 195-horseradish peroxidase-conjugated) (Boehringer Mannheim, F.R.G.), but using the described incubation buffer instead of that recommended by the manufacturer.

Only three patients showed a measurable TNFα concentration, two with septicemia (47 pg/ml and 720 pg/ml respectively) and one with a chronic autoimmune disorder (68 pg/ml).

To rule out the possibility that anti-TNFα antibodies might be complexed in vivo with TNFα, a C1q binding assay with immobilized C1q was performed according to the manufacturer's instructions (Bios, Munich, F.R.G.) using a polyclonal rabbit anti-TNFα antibody (1/500 in dilution buffer). As second antibody, the anti-rabbit IgG-alkaline phosphatase conjugate (1/1000) was used in dilution buffer.

Anti-TNFα antibodies were neither found in patients' sera nor in the immunoglobulin preparations using the ELISA procedure. Furthermore, the TNF/anti-TNFα immunocomplex assay showed no differences between the sera of patients and the negative controls.

In order to confirm the results obtained by ELISA, Western blot assays were performed with the human recombinant TNFα stabilized by human serum albumin. TNFα was boiled for 5 min in sample buffer with mercaptoethanol and SDS (Laemmli, 1970) (1600 U rTNFα in 20 µl sample buffer per lane) before application to SDS-PAGE (Laemmli, 1970). Proteins were transferred by semi-dry electroblotting to nitrocellulose (0.2 µm) (Schleicher & Schüll, Dassel, F.R.G.) in

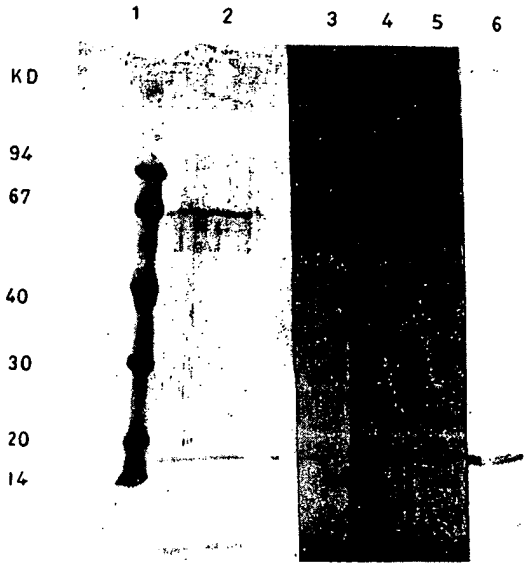


Fig. 1. Reaction of Venimmun and the sera of two patients with human rTNF and stabilization proteins in a Western blot blocked with Tween 20. Lane 1: standard proteins; lane 2: human rTNFα, both lanes stained with AuroDye; lane 3: Venimmun; lanes 4 and 5: sera of two patients; lane 6: polyclonal rabbit anti-human rTNFα (positive control) used in a parallel Western blot.

Tris/glycine/SDS/methanol buffer (48 mM Tris, 39 mM glycine, 0.0375% SDS, 20% methanol) for 1 h at 0.8 mA/cm^2. Blotted molecular weight markers (Pharmacia, Freiburg, F.R.G.) and the TNFα sample were stained using Auro Dye (Janssen, Beerse, Belgium). rTNFα was identified in monomeric form at the 17 kDa position (Fig. 1). Blocking and immunodetection were performed according to Fomsgaard et al. (1989) using Tween 20-Block. Patients' sera (diluted 1/5), the corresponding IgG fractions and the immunoglobulin preparations were applied at a concentration of 4 mg IgG/ml in blotting buffer (Fomsgaard et al., 1989) and incubated at room temperature overnight. The polyclonal rabbit anti-human rTNFα antibody was used as a positive control. The detector antibody, an HRPO-conjugated anti-human IgG and IgM antibody (Dianova, Hamburg, F.R.G.), was used at a dilution of 1/500 in incubation buffer. For the rabbit antibody an HRPO-conjugated anti-rabbit IgG antibody (Bio-Rad, Munich, F.R.G.) was used at a dilution of 1/500 in incubation buffer. After incubation at

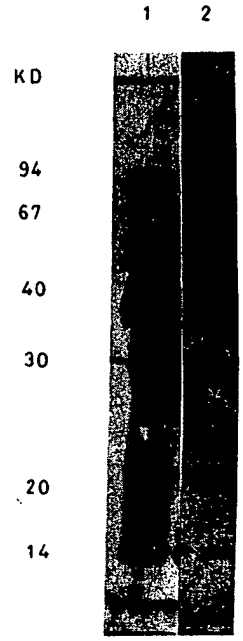


Fig. 2. Non-specific reaction of Venimmun with standard proteins in Western blot blocked with Tween 20. Lane 1: standard proteins, AuroDye protein stain; lane 2: Venimmun used as parallel Western blot.

room temperature for 2 h color was developed as described by Young et al. (1989). Bands were observed in the position of rTNFα (17 kDa) with all of the IgG fractions of patients' sera and the immunoglobulin preparations. Some of the diluted sera produced a band at 17 kDa together with a band at 68 kDa showing the position of the stabilization protein (Fig. 1). To confirm these results, we also used the non-fat dry milk block method according to the manufacturers instructions (Biozyme, Hameln, F.R.G.) because non-specific binding of immunoglobulins to protein on nitrocellulose when using detergent blocking (as done by Fomsgaard et al., 1989) has been documented (Batteiger et al., 1989). After blocking with non-fat dry milk no bands at 17 kDa or at the position of the stabilization proteins (68 kDa) were observed in the Western blot assay when analysing patients' sera or the immunoglobulin preparations, although a positive reaction was seen with the polyclonal anti-human rTNFα-immune serum in the 17 kDa position.

To demonstrate non-specific binding when using the Tween 20 detergent block, we tested the reaction of the patients sera and immunoglobulin preparations with the blotted standard proteins. In fact, under these experimental conditions all of the standard proteins showed non-specific binding with human IgG or IgM (Fig. 2).

In conclusion, using two immunochemical methods we failed to establish the presence of specific antibodies to TNFα in sera of human origin. Using the Western blot method described by Fomsgaard's group we found non-specific binding with the stabilization proteins and even with standard proteins. We cannot exclude the possibility that the binding of specific TNFα antibodies may be masked by the non-specific binding of immunoglobulins when using the Tween 20 block because Fomsgaard et al. (1989) were able to eliminate the reaction by absorption with soluble TNFα. Nevertheless, caution should be exercised when interpreting Western blot results after blocking with Tween 20. Furthermore, it will be necessary to consider which techniques should be used to assay naturally occurring specific anti-TNFα autoantibodies other than the inhibition of TNFα in biological systems.

References

Batteiger, B.E. (1988) Blocking of Immunoblots. Handbook of Immunoblotting of Proteins, Vol. 1. CRC Press, Boca Raton, FL, p. 145.

Bendtzen, K., Svenson, M., Fomsgaard, A. and Poulsen, L.K. (1989) Native inhibitors (autoantibodies) of IL-1α and TNF [letter]. Immunol. Today 10, 222.

Bendtzen, K., Svenson, M., Jønsson, V. and Hippe, E. (1990) Autoantibodies to cytokines – Friends or Foes? Immunol. Today 11, 167.

Foley, N., Lambert, C., McNicol, M., Johnson, N. and Rook, G.A.W. (1990) An inhibitor of the toxicity of tumor necrosis factor in the serum of patients with sarcoidosis, tuberculosis and Crohn's disease. Clin. Exp. Immunol. 80, 395.

Fomsgaard, A., Svenson, M. and Bendtzen, K. (1989) Autoantibodies to tumor necrosis factor α in healthy humans and patients with inflammatory diseases and gram-negative bacterial infections. Scand. J. Immunol. 30, 219.

Laemmli, U.I. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680.

Seckinger, P., Isaaz, S. and Dayer, J.-M. (1989) Purification and biologic characterization of a specific tumor necrosis factor α inhibitor. J. Biol. Chem. 20, 11966.

Young, P.R. (1989) An improved method for the detection of peroxidase-conjugated antibodies on immunoblots. J. Virol. Methods 24, 227.

Jonathan P. L. Cox[○], Ian M. Tomlinson[○■△] and Greg Winter[○■]

MRC Centre for Protein Engineering[○] and MRC Laboratory of Molecular Biology[■], Cambridge

A directory of human germ-line V_κ segments reveals a strong bias in their usage

From the genomic DNA of a single individual, we have amplified, cloned and sequenced 37 human germ-line V_κ segments. Four of these segments were new. We then compiled a comprehensive directory of all germ-line V_κ segments and identified 50 different sequences with open reading frames. Comparison with 236 rearranged sequences revealed that no more than 24 of these germ-line sequences could be assigned rearranged counterparts, that some of these were rarely used, and that only about 11 sequences are used frequently. This suggests that the expressed V_κ repertoire is mainly derived from a limited number of segments. Most surprisingly, the J_κ-distal region of the locus appears to be rarely used: we could unambiguously assign 162 rearranged sequences to V_κ segments of the J_κ-proximal region, but only 5 to segments of the J_κ-distal region.

1 Introduction

Recently antibodies of predefined binding specificity have been derived from repertoires of associated heavy and light chain variable (V) domains displayed on the surface of filamentous phage ([1]; for review see [2]). The process mimics production of antibodies by the immune system and bypasses hybridoma technology and immunization [3, 4]. Diverse repertoires of V domains have been provided by PCR amplification [5] of heavy and light chain V genes [6] from populations of lymphocytes. Alternatively the repertoires have been provided by *in vitro* rearrangement of cloned V segments [7]. To provide a bank of cloned V gene segments, we have amplified and sequenced V_H and V_λ segments from genomic DNA, leading to the isolation of almost all the functional human V_H segments [8] and many of the human V_λ segments [9]. Here we have attempted to gather a large selection of human germ-line V_κ segments using the same strategy. After submission of this report, a number of new V_κ segments were sequenced and mapped [10, 11]. For completeness we have included this information in our revised paper.

The human immunoglobulin light chain kappa (κ) locus is located on the short arm of chromosome 2 (2p11-12) [12, 13] and consists of a C_κ gene, 5 J_κ segments and at least 76 mapped germ-line V_κ segments [11, 14]. The V_κ segments may be divided by sequence homology [15] into three main subgroups, I–III, and several smaller subgroups (IV, V, VI and VII) [11, 16–18]. The portion of the locus harboring the majority of V_κ segments is thought to have arisen from a duplication event [19, 20]: 36 segments are located in the J_κ-distal region and 40 in the J_κ-proximal region [11, 14, 21, 22]. Segments are clustered in four distinct regions, A, B, L and O (Fig. 1) [18]. The A, L and O clusters are found within both the J_κ-distal and the J_κ-proximal regions, the B cluster only within the J_κ-proximal region.

Of the 76 mapped V_κ segments within the major locus, 57 have published sequences. Of these 57 segments, 48 have alleles with open reading frames (these correspond to 50 different sequences due to there being some segments from the J_κ-proximal and J_κ-distal regions which are identical and other segments which have multiple alleles) and 9 have frame shifts, stop codons or incomplete exons and are, therefore, regarded as pseudogenes. Most of the remaining 19 segments are also known to be pseudogenes [10, 16, 17, 19, 20, 23–36] (Fig. 1). There are also a number of orphon V_κ segments which have been mapped outside the major κ locus: W1-W11 [37, 38], V108 [39], Chr22-1 to Chr22-5 [40, 41], Chr1 [41], cos 118 [41] and Z1-Z4 [42, 43]. There are two further segments with open reading frames that have not been mapped (LFVK5 and LFVK431; L. Foroni, personal communication).

To clone the functional V_κ segments, we assembled a database of published germ-line sequences and then designed PCR primers to amplify each of the different subgroups. A germ-line V_κ segment consists (5′ to 3′) of a leader sequence (L), a leader intron, a continuation of the leader sequence (L′) and the V_κ exon (Fig. 2). Within the exon there are three framework regions (FR) and three complementarity-determining regions (CDR). Adjoining the 3′ end of the exon is a highly conserved heptamer recombination signal and a less conserved nonamer region, the two separated by 11–14 nucleotides. Foward PCR primers were based in the heptamer region and back primers on the overlap between the leader intron and L′ region (Fig. 2 and Table 1). Using these primers we amplified, cloned and sequenced germ-line V_κ segments from the genomic DNA extracted from the peripheral white blood cells of a single donor (DP).

2 Materials and methods

2.1 Primer design

Forward and back primers were designed for the V_κI, II, III, IV and VI subgroups from alignments of published germ-line sequences. For the V_κI subgroup, three cluster-specific

[I 12076]

△ Supported by the Medical Research Council Human Genome Mapping Project.

Correspondence: I. M. Tomlinson, MRC Centre for Protein Engineering, Hills Road, Cambridge CB2 2QH, GB (Fax: 02 23 40 21 40)

Abbreviations: **CDR:** Complementarity-determining region **FR:** Framework region

Key words: Human immunoglobulin kappa (V_κ) segments

0014-2980/94/0404-0827$10.00+.25/0

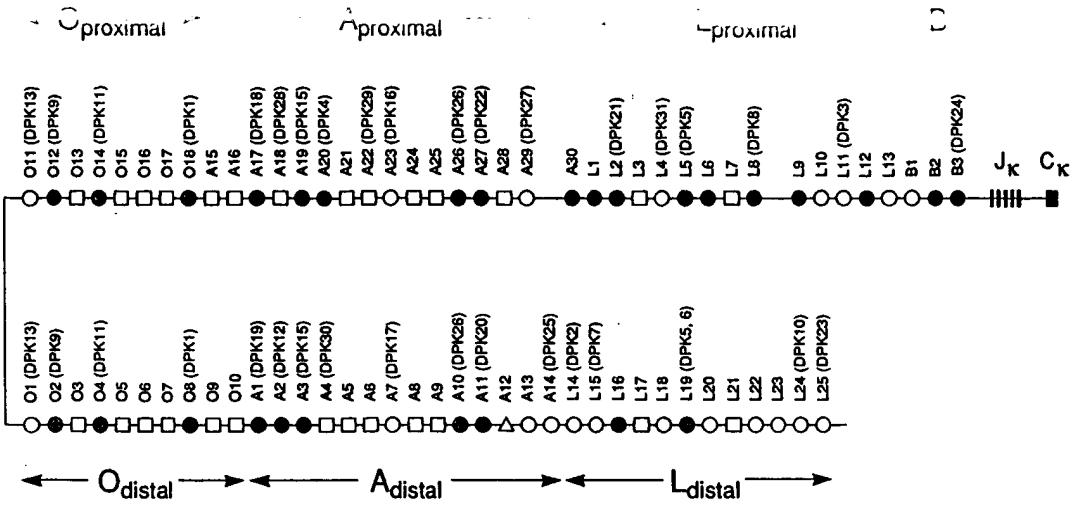


Figure 1. Schematic representation of the human immunoglobulin κ locus (after [11, 21]). Germ-line $V_κ$ segments with open reading frames are represented by circles. The segments A4, A5, A18, A22, A28, L3, L7, L17 and L21 have frame shifts, stop codons or incomplete $V_κ$ exons and are, therefore, regarded as pseudogenes (squares). The remaining 19 segments have either not been sequenced (triangle) or are known to be pseudogenes (squares). The segments L4 and L16 have a number of alleles, some with open reading frames and others with stop codons. Where a segment has been seen rearranged *in vivo*, the circle has been filled (black or shaded). If the rearranged sequences could not be unambigously assigned to segments of the $J_κ$-proximal or $J_κ$-distal regions, the corresponding segments from both regions have been shaded. Sequences from this study have been attributed their DPK codes (see text).

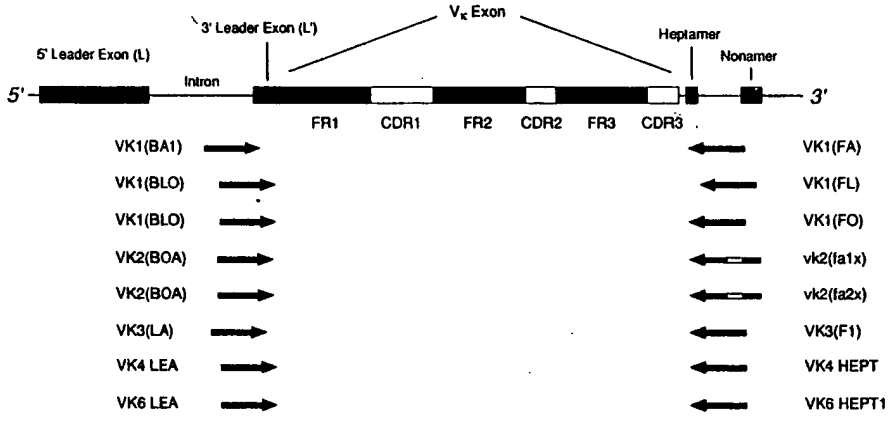


Figure 2. Location of PCR primers for amplification of the human germ-line $V_κ$ segment.

Table 1. Primers used for PCR amplification of the $V_κ$ exon

	Primer	Sequence (5' to 3')
$V_κI$	VK1(FA)	GCT CTA GAC GGG CTT GTA TCA CAG TG
	VK1(FL)	GCT CTA GAG TT(CT) (AG)GG T(GT)(GT) GTA ACA CT
	VK1(FO)	GCT CTA GAA TG(AC) CTT GT(TA) ACA CTG TG
	VK1(BA1)	CCC CCA AGC TTT GTT CCT AAT ATC AGA TA
	VK1(BLO)	CCC CCA AGC TTA ATC (TG)CA GGT (GT)CC AGA TG
$V_κII$	vk2(fa1x)	GAG GTT TTC TAG A(TG)G (GA)(GT)(CT) TGT A(GC)C ACT GTG
	vk2(fa2x)	GAG GTT TTC TAG AAG (GA)(GT)(CT) TGT A(GC)C ACT GTG
	VK2(BOA)	CCC CCA AGC TT(TA) A(TC)T TCA GGA TCC AGT G
$V_κIII$	VK3(F1)	GGA ATT CT(CT) A(TA)G (CT)TG AAT CAC TGT G
	VK3(LA)	CCC CCA AGC TTT CCA AT(TC) T(CT)(AG) GAT ACC AC
$V_κIV$	VK4 HEPT	GCT CTA GAC GAG GCT GAA GCA CTG TG
	VK4 LEA	CCC CCA AGC TTA CTA CAG GTG CCT ACG GG
$V_κVI$	VK6 HEPT1	GCT CTA GAG GGT TGT A(GA)C ACA GTG TG
	VK6 LEA	CCC CCA AGC TTT TTT CAG CCT CCA GGG GT

forward primers were designed, VK1(FA), VK1(FL), and VK1(FO), corresponding to the A, L and O clusters, respectively. Two cluster-specific back primers, VK1(BA1) and VK(BLO), corresponding to the A cluster and the combined B and L clusters were also used. *Hind*III restriction sites were incorporated into the 5′ end of all back primers. All forward primers had *Xba*I restriction sites at the 5′ end, except VK3(F1) which had an *EcoR*I site and the V_κII primers, where the *Xba*I site was introduced into the middle of the primer to replace a particularly degenerate section. The primers were used in pairs as indicated in Fig. 2.

2.2 PCR amplification and sequencing

PCR amplification was performed according to [8] with the following modifications. Products from the PCR amplifications and restriction enzyme digests were purified using Magic™ PCR Preps (Promega, Madison, WI). Inserts from plaques picked from a TYE [44] plate were amplified with an M13mp19-specific primer pair using the PCR (25 cycles, each cycle consisting of 1 min at 94 °C, 1 min at 55 °C, 30 s at 72 °C; at the end of 25 cycles there was a final extension at 65 °C for 5 min). Sequencing was performed using an M13-specific primer, Taq polymerase and fluorescent dideoxy chain terminators [45]. The sequences were analyzed on an Applied Biosystems 373A Automated DNA Sequencer (Foster City, CA).

3 Results

Amplification using the subgroup-specific primers with suitable adjustment of the annealing temperature gave single PCR bands of varying intensities. Primer pairs were found to be specific for their respective subgroups. We sequenced 142 clones (62 V_κI, 60 V_κII, 14 V_κIII, 1 V_κIV and 5 V_κVI). From these, 37 V_κ sequences were identified. Excluding those identical to mapped orphon segments (see below) 27 segments had open reading frames (DPK1-DPK27): 24 were identical to published sequences and 3 segments, DPK2, DPK14 and DPK23, were new. DPK2 was most similar to L1/HK137 (14 nucleotide changes), DPK14 to the pseudogene A5 (insertion of one nucleotide and 3 changes), and DPK23 to L10/Vh (8 nucleotide changes). The remaining segments are identical to the pseudogenes

Table 2. Assignment of rearranged V_κ sequences to their closest germline counterparts

	Germline sequence	No. of rearranged sequences[a]	Closest rearranged sequence Δ(N, P)[b]
V_κI	O8/O18/DPK1	21	(0, 0)
	A30	3	(5, 2)
	L1/HK137	1	(15, 8)
	A20/DPK4	9	(0, 0)
	L5/L19[1]/Vb/Vb′/V4b/DPK5	4	(3, 0)
	L8/Vd/DPK8	10	(1, 0)
	L9/Ve	3	(0, 0)
	L12[1]/HK102/V1	1	(37, 17)
	L12[2]	12	(5, 1)
	O12[1]/V3b	1	(2, 1)
	O2/O12[2]/DPK9	30	(0, 0)
	O4/O14/LFVK19H/DPK11	1	(11, 6)
		96	
V_κII	A2/DPK12	1	(9, 8)
	A3/A19/DPK15	11	(0, 0)
	A17/DPK18	11	(1, 0)
	A1/DPK19	1	(0, 0)
		24	
V_κIII	A11/humkv305/DPK20	1	(3, 2)
	L16/humkv328/humkv328h2	2	(3, 3)
	L2/humkv328h5/DPK21	20	(1, 0)
	A27/humkv325/Vκ RF/DPK22	48	(0, 0)
	L6/Vg	19	(0, 0)
		90	
V_κIV	B3/VkIV/DPK24	22	(1, 1)
V_κV	B2/EV15	2	(0, 0)
V_κVI	A10/A26/DPK26	2	(3, 1)

a) Number of rearranged sequences assigned to their closest germline counterpart.

b) Δ(N, P) denotes the number of nucleotide (N) and amino acid residue changes (P).

				FR1	CDR1	FR2	CDR2	FR3	CDR3
				1 10 20	30	40	50	60 70 80	90
$V_\kappa I$	7	R0	**O8/O18**[33]/DPK1	DIQMTQSPSSLSASVGDRVTITC	QASQDISN YLN	WYQQKPGKAPKLLIY	DASNLET	GVPSRFSGSGSGTDFTFTISSLQPEDIATYYC	QQYDNLP
	7		**L14**[10]/DPK2	NIQMTQSPSAMSASVGDRVTITC	RARQGISN YLA	WFQQKPGKVPKHLIY	AASSLQS	GVPSRFSGSGSGTEFTLTISSLQPEDFATYYC	LQHNSYP
	7		**L15**^{1}[10]/HK101[31]/HK146/HK189[19]	DIQMTQSPSSLSASVGDRVTITC	RARQGISS WLA	WYQQKPEKAPKSLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQYNSYP
	7		**L11**/Vf[33]/DPK3	AIQMTQSPSSLSASVGDRVTITC	RASQGIRN DLG	WYQQKPGKAPKLLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	LQDYNYP
	7	R2	**A30**[10]	DIQMTQSPSSLSASVGDRVTITC	RASQGIRN DLG	WYQQKPGKAPKRLIY	AASSLQS	GVPSRFSGSGSGTEFTLTISSLQPEDFATYYC	LQHNSYP
	7		LFVK5[*]	DIQMTQSPSSLSASVGDRVTITC	RASQGISN NLN	WYQQKPGKTPKLLIY	AAPSLQS	GIPSRFSDSGSGADYTLTIRSLQPEDFATYYC	QQSDSTP
	7		LFVK431[*]	DIQMTQSPSSLSASVGDRVTITC	RASQGISN YLA	WFQQKPGKAPKSLIY	AASSLQS	GVPSKFSGSGSGTDFTLTISSLQPEDFATYYC	QQYNSYP
	7	R8	**L1**[10]/HK137[19]	DIQMTQSPSSLSASVGDRVTITC	RASQGISN YLA	WFQQKPGKAPKSLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQYNSYP
	7	R0	**A20**[17]/DPK4	DIQMTQSPSSLSASVGDRVTITC	RASQGISN YLA	WYQQKPGKVPKLLIY	AASTLQS	GVPSRFSGSGSGTDFTLTISSLQPEDVATYYC	QKYNSAP
	7		**L18**^{1}[10]/Va''[20]	AIQLTQSPSSLSASVGDRVTITC	RASQGISS ALA	WYQQKPGKAPKLLIY	DASSLES	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQFNNYP
	7		**L4/L18**^{2}/Va'/V4a[35]	AIQLTQSPSSLSASVGDRVTITC	RASQGISS ALA	WYQQKPGKAPKLLIY	DASSLES	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQFNSYP
	7	R0	**L5/L19**^{1}[10]/Vb/Vb'/V4b[35]/DPK5	DIQMTQSPSSVSASVGDRVTITC	RASQGISS WLA	WYQQKPGKAPKLLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQANSFP
	7		**L19**^{2}/Vb''[35]/DPK6	DIQMTQSPSSVSASVGDRVTITC	RASQGISS WLA	WYQQKPGKAPKLLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQANSFP
	7		**L15**^{2}[10]/HK134/HK166[19]/DPK7	DIQMTQSPSSLSASVGDRVTITC	RASQGISS WLA	WYQQKPEKAPKSLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQYNSYP
	7	R0	**L8**/Vd[35]/DPK8	DIQLTQSPSFLSASVGDRVTITC	RASQGISS YLA	WYQQKPGKAPKLLIY	AASTLQS	GVPSRFSGSGSGTEFTLTISSLQPEDFATYYC	QQLNSYP
	7	R0	**L9**/Ve[35]	AIRMTQSPSSFSASTGDRVTITC	RASQGISS YLA	WYQQKPGKAPKLLIY	AASTLQS	GVPSRFSGSGSGTDFTLTISCLQSEDFATYYC	QQYYSYP
	7	R17	**L12**^{1}/HK102[31]/V1[25]	DIQMTQSPSTLSASVGDRVTITC	RASQSISS WLA	WYQQKPGKAPKLLIY	DASSLES	GVPSRFSGSGSGTEFTLTISSLQPDDFATYYC	QQYNSYS
	7	R1	**L12**^{2}[10]	DIQMTQSPSTLSASVGDRVTITC	RASQSISS WLA	WYQQKPGKAPKLLIY	KASSLES	GVPSRFSGSGSGTEFTLTISSLQPDDFATYYC	QQYNSYS
	7	R1	**O12**^{1}/V3b[29]	DIQMTQSPSPLSASVGDRVTITC	RASQSISS YLN	WYQQKPGKAPKLLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQGYSTP
	7	R0	**O2/O12**^{2}[29]/DPK9	DIQMTQSPSSLSASVGDRVTITC	RASQSISS YLN	WYQQKPGKAPKLLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQSYSTP
	7		**L24**[10]/Ve*/V13[25]/DPK10	VIWMTQSPSLLSASTGDRVTISC	RMSQGISS YLA	WYQQKPGKAPELLIY	AASTLQS	GVPSRFSGSGSGTDFTLTISCLQSEDFATYYC	QQYYSFP
	7	R6	**O4/O14**[29]/LFVK19R[*]/DPK11	DIQLTQSPSSLSASVGDRVTITC	RVSQGISS YLN	WYRQKPGKVPKLLIY	SASNLQS	GVPSRFSGSGSGTDFTLTISSLQPEDVATYYG	QRTYNAP
	7		**L22**[10]	DIQMTQSPSFLSASVGDRVSIIC	WASEGISS NLA	WYLQKPGKSPKLFLY	DAKDLHP	GVSSRFSGRGSGTDFTLTIISLKPEDFAAYYC	KQDFSYP
	7		**L23**[10]	AIRMTQSPFSLSASVGDRVTITC	WASQGISS YLA	WYQQKPAKAPKLFIY	YASSLQS	GVPSRFSGSGSGTDYTLTISSLQPEDFATYYC	QQYYSTP
$V_\kappa II$	12	R8	**A2**[23]/DPK12	DIVMTQTPLSLSVTPGQPASISC	KSSQSLLHS DGKTYLY	WYLQKPGQPPQLLIY	EVSNRFS	GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC	MQSIQLP
	13		**O1/O11**^{1}[29]/DPK13	DIVMTQTPLSLPVTPGEPASISC	RSSQSLLDSDDGNTYLD	WYLQKPGQSPQLLIY	TLSYRAS	GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC	MQRIEFP
	13		**O11**^{2}/V3a[29]	ASISC	RSSQSLLDSDDGNTYLD	CYLQKPGQSPQLLIY	TLSYRAS	GVPDRFSDSGSGTDFTLKISRVEAEDVGVYYC	MQRIEFP
	12		**L13**[10]	DIVMTQHLLSLPIPLGEPASISC	RSSQSLLHS DGNTYLD	WYLQKPGQSPQLLIY	TISNKFY	GVPNKFSGSRSGTGFTLKFSKVEAEDVGVYCC	EQGLQGP
	12		DPK14	EIVMTQTPLSLSITPGEQASMSC	RSSQSLLHS DGYTYLY	WFLQKARPVSTLLIC	EVSNRFS	GVPDRFSGSGSGTDFTLKISRVEAEDFGVYYC	MQDAQDP
	12	R0	**A3**[24]/**A19**[17]/DPK15	DIVMTQSPLSLPVTPGEPASISC	RSSQSLLHS NGYNYLD	WYLQKPGQSPQLLIY	LGSNRAS	GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC	MQALQTP
	12		**A29**[17]/DPK27	DIVMTQTPPSLPVNPGEPASISC	RSSQSLLHS NGYTYLH	WYLQKPGQSPQLLIY	RVSNHLS	GVPDRFSGSGSGSDFTLKISWVEAEDVGVYYC	MQATQFP
	12		**A13**[17]	DIVMTQTPPSLPVNPGEPASISC	RSSQSLLHS NGYTYLH	WYPQKPGQSPQLLIY	RVSSRFS	GVPDRFSGSGSGSDFTLKISWVEAEDVGVYYC	MQATQFP
	12		**A23**[24]/DPK16	DIVMTQTPLSSPVTLGQPASISC	RSSQSLVHS DGNTYLS	WLQQRPGQPPRLLIY	KISNRFS	GVPDRFSGSGAGTDFTLKISRVEAEDVGVYYC	MQATQFP
	12		**A7**[17]/DPK17	DIVMTQTPLSSPVTLGQPASISF	RSSQSLVHS DGNTYLS	WLQQRPGQPPRLLIY	KVSNRFS	GVPDRFSGSGAGTDFTLKISRVEAEDVGVYYC	TQATQFP
	12	R0	**A17**[17]/DPK18	DVVMTQSPLSLPVTLGQPASISC	RSSQSLVYS DGNTYLN	WFQQRPGQSPRRLIY	KVSNRDS	GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC	MQGTHWP
	12	R0	**A1**[17]/DPK19	DVVMTQSPLSLPVTLGQPASISC	RSSQSLVYS DGNTYLN	WFQQRPGQSPRRLIY	KVSNWDS	GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC	MQGTHWP
$V_\kappa III$	8	R2	**A11**[24]/humkv305[27]/DPK20	EIVLTQSPATLSLSPGERATLSC	GASQSVSSS YLA	WYQQKPGLAPRLLIY	DASSRAT	GIPDRFSGSGSGTDFTLTISRLEPEDFAVYYC	QQYGSSP
	7		**L20**/Vg*[20]	EIVLTQSPATLSLSPGERATLSC	RASQGVSS YLA	WYQQKPGQAPRLLIY	DASNRAT	GIPARFSGSGPGTDFTLTISSLEPEDFAVYYC	QQRSNWH
	7	R3	**L16**[10]/humkv328/humkv328h2[34]	EIVMTQSPATLSVSPGERATLSC	RASQSVSS NLA	WYQQKPGQAPRLLIY	GASTRAT	GIPARFSGSGSGTEFTLTISSLQSEDFAVYYC	QQYNNWP
	7	R0	**L2**[10]/humkv328h5[34]/DPK21	EIVMTQSPATLSVSPGERATLSC	RASQSVSS NLA	WYQQKPGQAPRLLIY	GASTRAT	GIPARFSGSGSGTEFTLTISSLQSEDFAVYYC	QQYNNWP
	8	R0	**A27**[24]/humkv325/Vk RF[28]/DPK22	EIVLTQSPGTLSLSPGERATLSC	RASQSVSSS YLA	WYQQKPGQAPRLLIY	GASSRAT	GIPDRFSGSGSGTDFTLTISRLEPEDFAVYYC	QQYGSSP
	8		**L25**[10]/DPK23	EIVMTQSPATLSLSPGERATLSC	RASQSVSSS YLS	WYQQKPGQAPRLLIY	GASTRAT	GIPARFSGSGSGTDFTLTISSLQPEDFAVYYC	QQDYNLP
	8		**L10**^{1}[10]	EIVMTQSPPTLSLSPGERVTLSC	RASQSVSSS YLS	WYQQKPGQAPRLLIY	GASTRAT	GIPARFSGSGSGTDFTLTISSLQPEDFAVYYC	QQDYNLP
	8		**L10**^{2}/Vh[32]	EIVMTQSPPTLSLSPGERVTLSC	RASQSVSSS YLT	WYQQKPGQAPRLLIY	GASTRAT	SIPARFSGSGSGTDFTLTISSLQPEDFAVYYC	QQDHNLP
	7	R0	**L6**/Vg[32]	EIVLTQSPATLSLSPGERATLSC	RASQSVSS YLA	WYQQKPGQAPRLLIY	DASNRAT	GIPARFSGSGSGTDFTLTISSLEPEDFAVYYC	QQRSNWP
$V_\kappa IV$	13	R1	**B3**/VkIV[16]/DPK24	DIVMTQSPDSLAVSLGERATINC	KSSQSVLYSSNNKNYLA	WYQQKPGQPPKLLIY	WASTRES	GVPDRFSGSGSGTDFTLTISSLQAEDVAVYYC	QQYYSTP
$V_\kappa V$	7	R0	**B2**/EV15[30]	ETTLTQSPAFMSATPGDKVNISC	KASQDIDD DMN	WYQQKPGEAAIFIIQ	EATTLVP	GIPPRFSGSGYGTDFTLTINNIESEDAAYYFC	LQHDNFP
$V_\kappa VI$	7		**A14**[26]/DPK25	DVVMTQSPAFLSVTPGEKVTITC	QASEGIGN YLY	WYQQKPDQAPKLLIK	YASQSIS	GVPSRFSGSGSGTDFTFTISSLEAEDAATYYC	QQGNKHP
	7	R1	**A10**[26]/**A26**[17]/DPK26	EIVLTQSPDFQSVTPKEKVTITC	RASQSIGS SLH	WYQQKPDQSPKLLIK	YASQSFS	GVPSKFSGSGSGTDFTLTINSLEAEDAATYYC	HQSSSLP
$V_\kappa VII$	11		**B1**[36]	DIVLTQSPASLAVSPGQRATITC	RASESVSFL GINLIH	WYQQKPGQPPKLLIY	QASNKDT	GVPARFSGSGSGTDFTLTINPVEANDTANYYC	LQSKNFP

Figure 3. Amino acid sequences of germ-line V_κ segments with open reading frames. Where possible, sequences have been assigned to their respective loci (bold type). Where two alleles have different nucleotide sequences these are followed by the superscripts ^{1} or ^{2}. Amino acid sequences from this study are labeled DPK1-DPK27 (see text). References for other sequences are in brackets, [*] indicates L. Foroni, personal communication. Amino acids are shown in single-letter codes. Sequences are arranged alphabetically within each subgroup according to the sequence of CDR1. Where two CDR1 sequences are identical, the order is based similarly upon the sequence of CDR2 or CDR3. The hypervariable regions, CDR1, CDR2 and CDR3 as defined by Kabat et al. [15] are labeled, as are the antigen-binding loops, L1, L2 and L3 as defined by Chothia and Lesk [50]. Numbering of amino acid residues is according to [15] except in CDR1 where numbering is according to [50]. The length of the L1 loop [50] is shown in italics. Germ-line sequences with known rearranged counterparts have been labeled R*N*, where *N* represents the number of amino acid differences between the germ-line V_κ sequence and its closest rearranged counterpart (see Table 2). Note that L5/L19^1/Vb/Vb'/V4b/DPK5 and L19^2/Vb''/DPK6; and L16/humkv328/humkv328h2 and L2/humkv328h5/DPK21 have identical protein sequences but different nucleotide sequences (see Fig. 4).

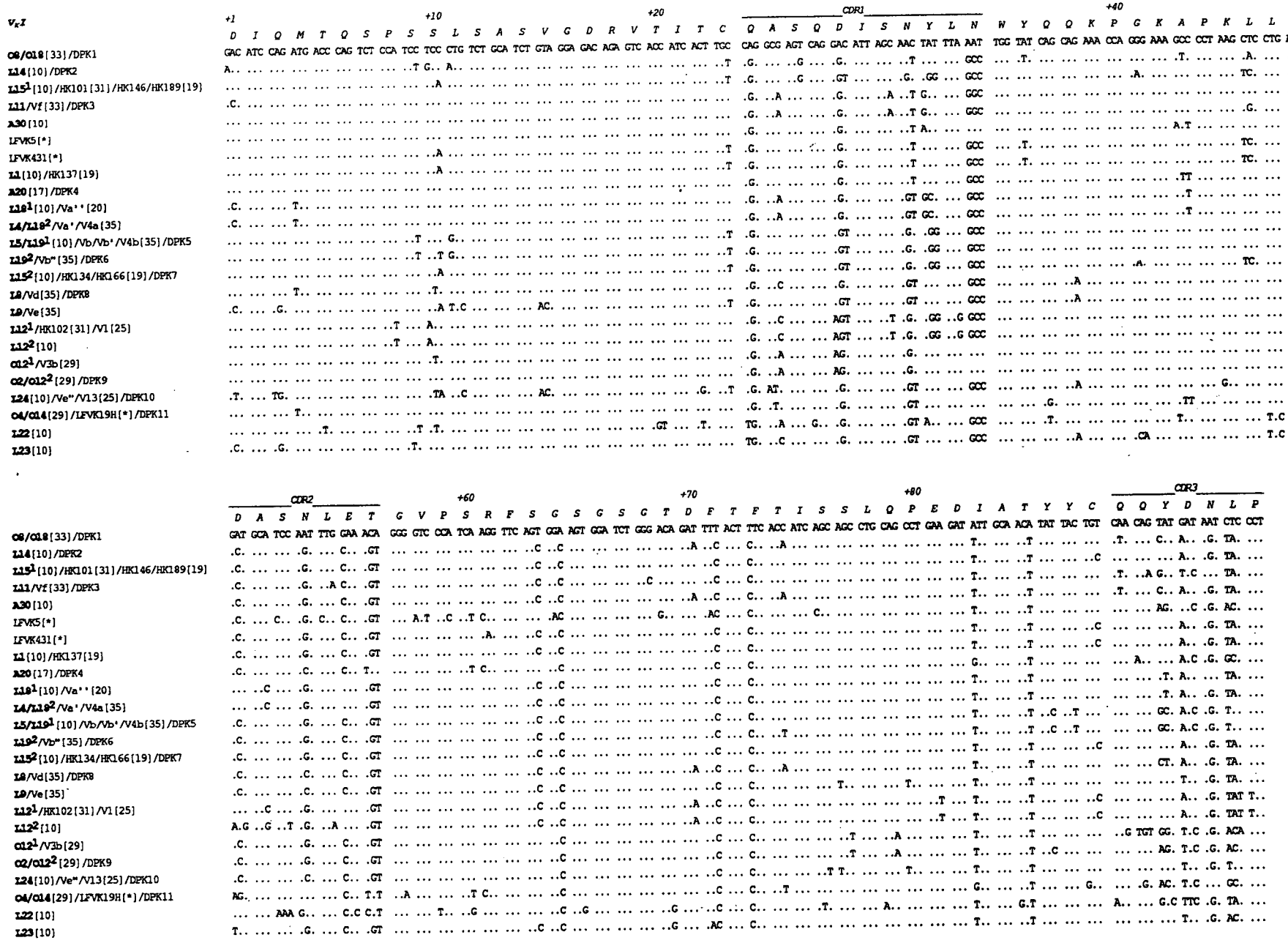
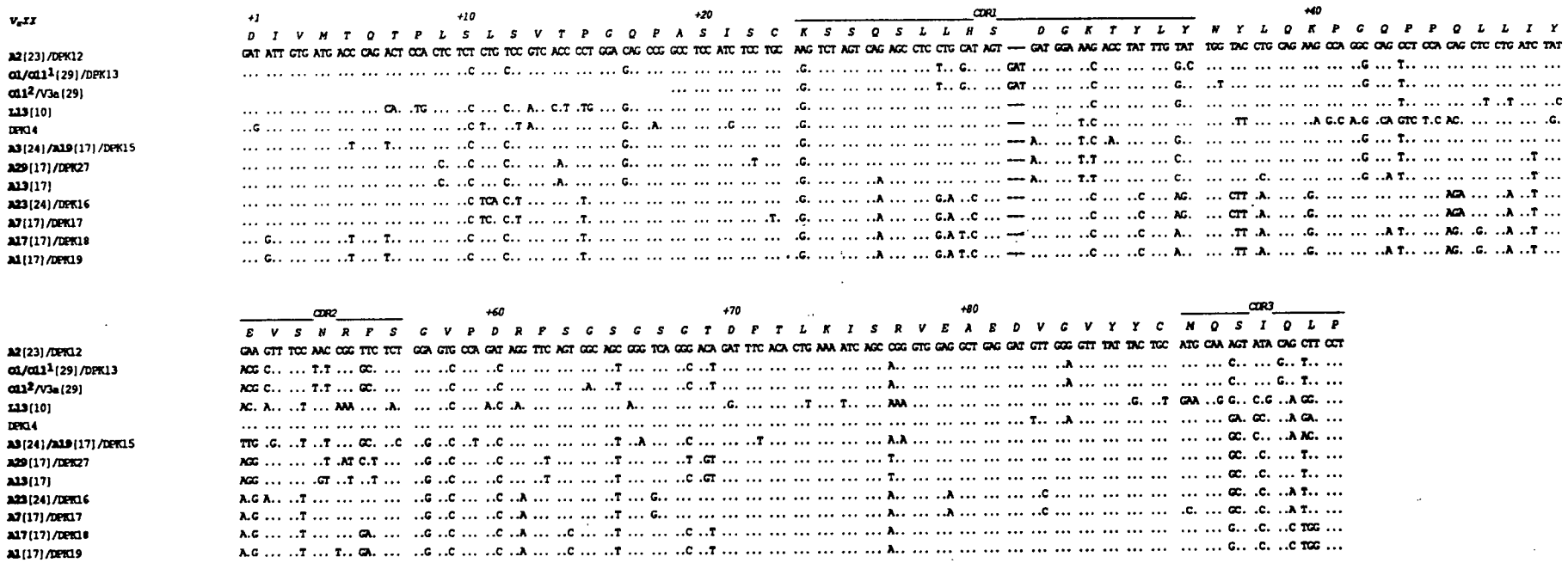

Vκ I
+1 +10 +20 CDR1 +40
D I Q M T Q S P S S L S A S V G D R V T I T C Q A S Q D I S N Y L N W Y Q Q K P G K A P K L L I Y
O8/O18[33]/DPK1
L14[10]/DPK2
L15[1][10]/HK101[31]/HK146/HK189[19]
L11/Vf[33]/DPK3
A30[10]
LFVK5[*]
LFVK431[*]
L1[10]/HK137[19]
A20[17]/DPK4
L18[1][10]/Va''[20]
L4/L18[2]/Va'/V4a[35]
L5/L19[1][10]/Vb/Vb'/V4b[35]/DPK5
L19[2]/Vb"[35]/DPK6
L15[2][10]/HK134/HK166[19]/DPK7
L8/Vd[35]/DPK8
L9/Ve[35]
L12[1]/HK102[31]/V1[25]
L12[2][10]
O12[1]/V3b[29]
O2/O12[2][29]/DPK9
L24[10]/Ve"/V13[25]/DPK10
O4/O14[29]/LFVK19H[*]/DPK11
L22[10]
L23[10]
CDR2 +60 +70 +80 CDR3
D A S N L E T G V P S R F S G S G S G T D F T F T I S S L Q P E D I A T Y Y C Q Q Y D N L P

Figure 4


Vκ II
+1 +10 +20 CDR1 +40
D I V M T Q T P L S L S V T P G Q P A S I S C K S S Q S L L H S — D G K T Y L Y W Y L Q K P G Q P P Q L L I Y
A2[23]/DPK12
O1/O11[29]/DPK13
O11[2]/V3a[29]
L13[10]
DPK14
A3[24]/A19[17]/DPK15
A29[17]/DPK27
A13[17]
A23[24]/DPK16
A7[17]/DPK17
A17[17]/DPK18
A1[17]/DPK19
CDR2 +60 +70 +80 CDR3
E V S N R F S G V P D R F S G S G S G T D F T L K I S R V E A E D V G V Y Y C N Q S I Q L P
A2[23]/DPK12
O1/O11[29]/DPK13
O11[2]/V3a[29]
L13[10]
DPK14
A3[24]/A19[17]/DPK15
A29[17]/DPK27
A13[17]
A23[24]/DPK16
A7[17]/DPK17
A17[17]/DPK18
A1[17]/DPK19
Vκ III
+1 +10 +20 CDR1 +40
E I V L T Q S P A T L S L S P G E R A T L S C G A S Q S V S S S Y L A W Y Q Q K P G L A P R L L I Y
A11[24]/humkv305[27]/DPK20
L20/Vg-[20]
L16[10]/humkv328/humkv328h2[34]
L2[10]/humkv328h5[34]/DPK21
A27[24]/humkv325/Vκ RF[28]/DPK22
L25[10]/DPK23
L10[1][10]
L10[2]/Vh[32]
L6/Vg[32]
CDR2 +60 +70 +80 CDR3
D A S S R A T G I P D R F S G S G S G T D F T L T I S R L E P E D F A V Y Y C Q Q Y G S S P
A11[24]/humkv305[27]/DPK20
L20/Vg-[20]
L16[10]/humkv328/humkv328h2[34]
L2[10]/humkv328h5[34]/DPK21
A27[24]/humkv325/Vκ RF[28]/DPK22
L25[10]/DPK23
L10[1][10]
L10[2]/Vh[32]
L6/Vg[32]

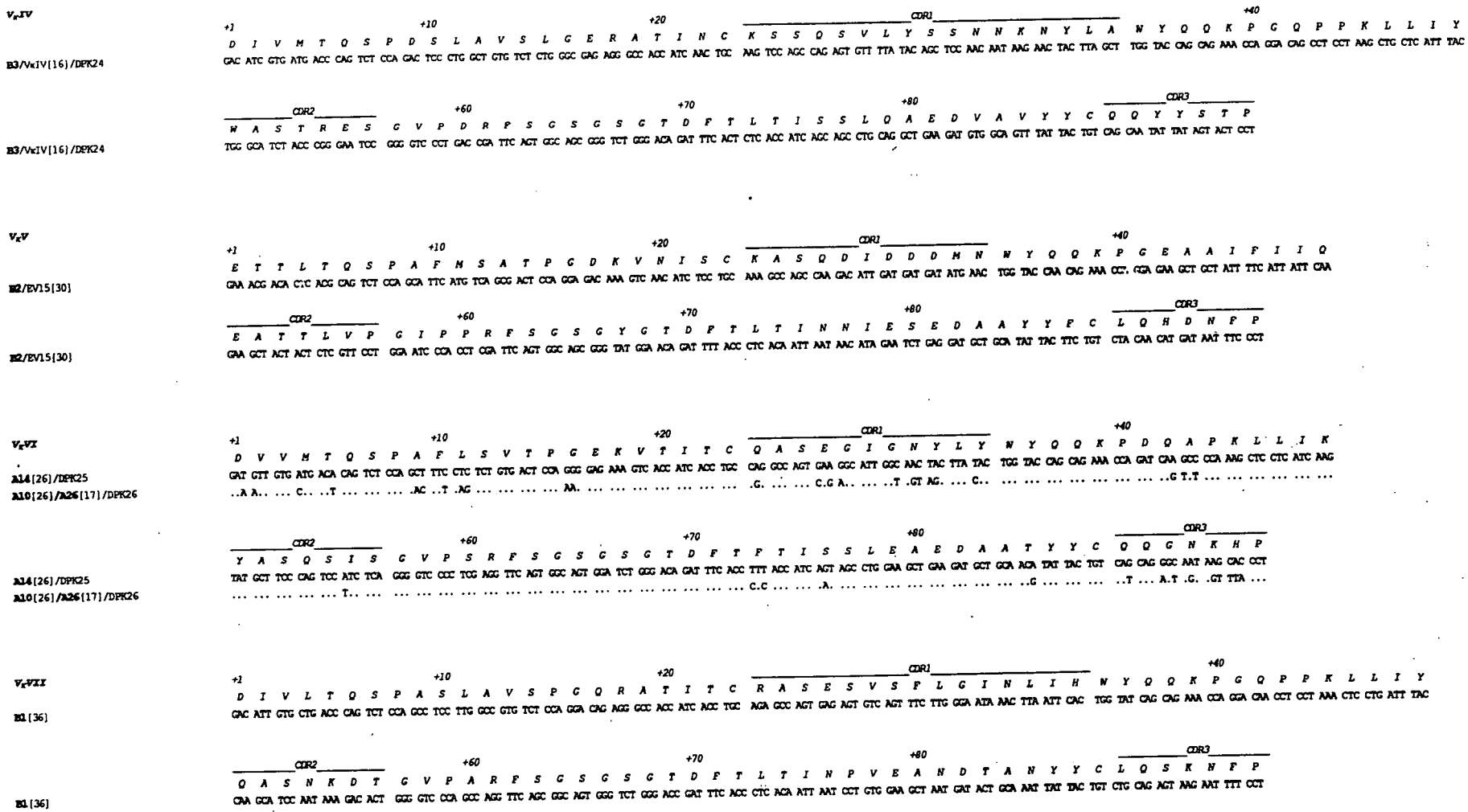


Figure 4. Nucleotide sequences of germ-line V_x segments with open reading frames. Sequences were extracted from the Genbank/EMBL nucleotide databases wherever possible. Sequences from this study have been labeled with their DPK codes (see text). References for other sequences are in brackets. [*] denotes L. Foroni, personal communication. A translation Sequences from this study have been labeled with their DPK codes (see text). References for other sequences are in brackets. [*] denotes L. Foroni, personal communication. A translation of the master sequence is shown directly above it: residues are numbered according to [15]. Nucleotides identical to the master sequence are represented by a dot and insertions are represented by a dash. The directory of human germ-line V_x sequences is available on disk on request.

A4 [25], A18 [27], A23 [24], L4 (Va) [35], HK100 [31] and to the orphon segments W2 [38], W8/W10 [38], Z2 [43], chr22-4 [41]: DPK37 is new and is most similar to the orphon Z3 [41] (deletion of one nucleotide and three changes). The sequences of the four new segments, DPK2, DPK14, DPK23 and DPK37, were confirmed with clones from independent PCR amplifications. Subsequently, two of these segments, DPK2 and DPK23, proved to be identical to the sequences L14 and L25, which were published after the submission of this paper [10]. The fact that 35 of the 37 segments amplified from DP are identical to published sequences confirms that germ-line V_κ sequence polymorphism is rather limited [46].

In addition to the confirmed sequences, many clones had sequences which were not observed in independent amplifications. Some had one or two nucleotide differences from the V_κ segments DPK1-DPK37 and were probably due to PCR errors. The 21 nucleotide substitutions in the 52 V_κI sequences correspond to 5×10^{-5} changes per nucleotide per PCR cycle and are consistent with the error rate for Taq polymerase [47]. Other clones appeared to be due to PCR cross-over [8, 48].

We have compiled a comprehensive directory of the amino acid and nucleotide sequences of all germ-line V_κ segments with open reading frames (Figs. 1, 3 and 4). Where possible, sequences have been assigned to their respective loci. We have compared these sequences with those of 113 rearranged V_κ sequences taken from the Genbank/EMBL nucleotide databases and 123 additional sequences from the literature (our rearranged database is available on request). All rearranged sequences were assigned to their closest germ-line counterparts. In Table 2 we note the closest sequence identities and the number of rearranged sequences assigned to each germ-line sequence. In a few examples, the rearranged sequence was found to be a composite of two V_κ segments, presumably arising by PCR cross-over during cloning. For example, V_κI-X14 [49] is identical to O4/O14/LFVK19H/DPK11 over the first 220 nucleotides and virtually identical to O2/O12[2]/DPK9 over the last 100 nucleotides, due to a cross-over in FR3.

Only 24 of the 50 germ-line sequences in our directory were found to have rearranged counterparts. Although 162 rearranged sequences correspond to mapped V_κ segments from the J_κ-proximal region of the kappa locus only 5 are derived from the J_κ-distal region. These correspond to the germ-line sequences A2/DPK12, A1/DPK19, A11/humkv-305/DPK20 and L16/humkv328/humkv328h2 (see Fig. 1). Of the rearranged sequences, 69 correspond to O8/O18/-DPK1, L5/L19/Vb/Vb'/V4b/DPK5, O2/O12[2]/DPK9, O4/-O14/LFVK19H, A3/A19/DPK15 and A10/A26/DPK26: in these cases each rearranged sequence could have been derived from V_κ segments from either the J_κ-proximal or J_κ-distal region. No rearranged sequences correspond to the unmapped V_κ segments LFVK5, LFVK431 and DPK14.

The distribution of the number of amino acid differences across all 236 assigned rearranged sequences is shown in Fig. 5. This includes sequences of antibodies with a wide range of specificities. As sequence polymorphism at the germ-line level is limited (see above) the majority of changes are probably due to somatic mutation.

4 Discussion

4.1 Functional V_κ segments

The map of the V_κ locus (Fig. 1) contains 76 segments. It is estimated, on the basis of hybridization experiments, that 4–6 V_κ segments have yet to be mapped [11, 14]. The V_κ segments LFVK5, LFVK431 and DPK14 could, therefore, correspond to the unmapped loci or could be allelic variants of mapped V_κ segments.

From our sequence directory (Fig. 3 and 4) and from the assignment of 236 rearranged sequences to their closest germ-line counterparts (Table 2) we find that no more than 24 of the 50 germ-line sequences have rearranged counterparts, suggesting that the expressed V_κ repertoire is mainly derived from these segments. Indeed for 7 of these V_κ sequences we could find only a single example of a rearrangement, and in 4 cases (L1/HK137, L12[1]/HK102/V1, O4/O14/LFVK19H/DPK11 and A2/DPK11) the assignment is tentative in view of the number of sequence differences. It is also possible that some of the remaining segments with open reading frames are functional, but are only rarely used. Others may well prove to have frame shifts or stop codons in the leader exon, defective recombination signals or splice sites or other *cis*-acting defects which render them non functional. A14/DPK25, for example, has an altered heptamer which may prevent its rearrangement. We have made similar observations for human germ-line V_H segments where only 49/83 germ-line sequences with open reading frames were seen as rearranged genes [8].

4.2 V_κ segment usage

The assignment of rearranged V_κ genes to their germ-line counterparts (Table 2) confirms that segments from all four clusters (A, B, L and O) and from six subgroups (I–VI) are used. Segments from the V_κI and V_κIII subgroups are used most frequently (96 and 90 rearranged sequences respectively); those from the V_κII and V_κIV subgroups less frequently (24 and 22 respectively); and those from the V_κV and V_κVI subgroups rarely (2 rearranged sequences each). The only V_κVII segment, B1, has no rearranged counterpart in our database. As noted above, some of the V_κ

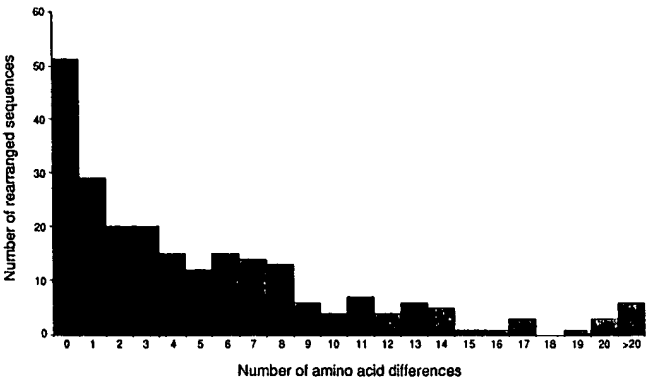


Figure 5. Distribution of the number of amino acid differences between each sequence in a database of rearranged V_κ sequences (236 in total) and its closest germ-line counterpart in our directory. The database of rearranged V_κ sequences is available on disk on request.

segments are rarely used, and of the 24 germ-line sequences, only about 11 appear to be frequently used (Table 2).

Of the rearranged sequences, 162 correspond to V_κ segments from the J_κ-proximal region of the kappa locus, whereas only 5 appear to be derived from the J_κ-distal region. Indeed, one of these rearranged sequences from the J_κ-distal region, EVJK11, which corresponds to A11/humkv305/DPK20, is the product of aberrant recombination [30]. An additional 69 rearranged sequences could conceivably be derived from V_κ segments in either the proximal or distal regions. However, as has been previously noted, the segments O2 (distal) and O12 (proximal) differ by one nucleotide in a region which flanks the exon [29]: rearranged sequences which include this region appear to be derived from O12 rather than O2 [29]. Although there is a clear bias towards use of the J_κ-proximal V_κ segments, we can detect no bias within the J_κ-proximal region.

The reason for the remarkable bias towards use of V_κ segments from the J_κ-proximal region is unclear. It may be due to the relative distances of the proximal and distal regions from the J_κ segments or the different recombination mechanisms necessary to produce the rearranged V_κ gene: most V_κ segments in the proximal region rearrange by deletion, whereas those in the distal region must rearrange by inversion. We note that lack of the J_κ-distal copy of the locus (haplotype 11) does not appear to be deleterious [21].

The assignment of rearranged V_κ genes may help in dissecting the mechanisms of the human immune system. For example, more than half of the rearranged sequences in our database have three or fewer residue changes (Fig. 5). Antibodies from patients with chronic lymphocytic leukemia (CLL) and X-linked agammaglobulinemia (XLA) are rarely mutated or unmutated. More highly mutated V_κ genes tend to be seen in myelomas and in antibodies which are subject to an antigen-driven response. In the case of autoantibodies, the level of somatic mutation appears to be normal.

4.3 Structures of loops implicit in the human V_κ segments

The antigen-binding loops of immunoglobulin variable domains have been shown to adopt a limited number of main chain conformations or "canonical structures" [50, 51]. The structure of each loop depends on its length and the identity of certain key residues involved in its packing, and using this information it is sometimes possible to predict the structure of the loops from the sequence of the V_κ domain [51]. Here we have attempted to identify the loop structures of those human V_κ segments that appear to undergo rearrangement (Table 2).

Across different species, the V_κ antigen binding site loop L1 (residues 26–32, corresponding to CDR1, see Fig. 3) can have lengths of 6, 7, 8, 11, 12 and 13 residues and presumably can form at least six major conformations [50, 51]. The structures of antibodies with loops of 6, 7, 12 and 13 residues have been solved crystallographically: in each case residue 29 in the loop is buried in the β-sheet framework, and packs against residues 2, 25, 33 and 71 [50, 51]. These packing contacts are often very similar: Ile/Val 2, Ala/Ser 25, Val/Ile/Leu 29, Leu/Met 33 and Tyr/Phe 71 [50].

However, for those human V_κ segments with rearranged counterparts (Table 2), we only see L1 loops of 7, 8, 12 and 13 residues (Fig. 3). In all cases the packing contacts are highly conserved, suggesting that the segments should encode four major conformations of the L1 loop. Of the germ-line sequences, 17 (141 rearranged sequences) encode a 7-residue L1 loop, 2 sequences (49 rearranged sequences of which 48 use the V_κIII segment A27) encode an 8-residue loop, 4 sequences (24 rearranged sequences) encode a 12-residue loop and 1 sequence (22 rearranged sequences) encodes a 13-residue L1 loop. L1 lengths of 6 residues have only been seen in mice [15], suggesting that mouse V_κ segments encode structures which cannot be encoded by human V_κ segments.

The L2 loop (residues 50–52, corresponding to CDR2, see Fig. 3) is only three residues long in all structures, and undergoes packing interactions with framework residues 48 and 64 [50, 51]. All germ-line V_κ sequences with rearranged counterparts (Fig. 3 and Table 2) have Ile 48 and Gly 64, indicating that they are likely to encode a single loop conformation.

The L3 loop (residues 91–96, corresponding to CDR3, see Fig. 3) is encoded mainly by the V_κ segment, and is most often a six-residue loop with Gln, Asn or His at 90 and Pro at 95 [50, 51]. Almost all human germ-line V_κ sequences with rearranged counterparts encode Gln 90 and Pro 95, except A20/DPK4 (Lys 90), O4/O14/LFVK19H/DPK11 (Arg 90), $L12^1$/HK102/V1 and $L12^2$ (Ser 95). This indicates that most human V_κ germ-line segments encode a single conformation of this loop: but depending on the location of the V-J join and nucleotide addition, other conformations may be formed in the rearranged gene.

We note that several of the germ-line sequences that do not have rearranged counterparts in our database (Fig. 3 and Table 2) have atypical residues at the key residues involved in the packing of the antigen-binding loops (L1 loop: L22 [Leu 48]; LFVK5 and $O11^2$/V3a [Asp 64]; L3 loop: L20/Vg" [His 95]) or unusual loop lengths (B1 has an L1 of 11 residues).

In combination, the three antigen-binding loops L1, L2 and L3 of human germ-line V_κ segments are likely to encode four major folds. Since our repertoire of cloned human germ-line V_κ segments includes 15 of the 24 germ-line V_κ sequences with rearranged counterparts and examples of each of the four major folds it should be a valuable resource for building synthetic antibodies for use in phage display libraries.

We would like to thank D. J. Perry for numerous donations of his blood, M. Llewelyn, C. Brown, G. Walter, P. Dear and G. Cook for their help. We would like to thank H. G. Zachau for providing L-region sequence data after first submission of our paper and L. Foroni for unpublished sequences. We would also like to thank C. Chothia for his comments. EMBL Data Library accession numbers for the sequences DPK2, DPK14, DPK23 and DPK37 are Z27498–Z27501.

Received July 14, 1993; in final revised form December 21, 1993; accepted December 23, 1993.

5 References

1 McCafferty, J., Griffiths, A. D., Winter, G. and Chiswell, D. J., *Nature* 1990. *348:* 552.
2 Hoogenboom, H. R., Marks, J. D., Griffiths, A. D. and Winter, G., *Immunol. Rev.* 1992. *130:* 41.
3 Clackson, T., Hoogenboom, H. R., Griffiths, A. D. and Winter, G., *Nature* 1991. *352:* 624.
4 Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D. and Winter, G., *J. Mol. Biol.* 1991. *222:* 581.
5 Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and Arnheim, N., *Science* 1985. *230:* 1350.
6 Orlandi, R., Gussow, D. H., Jones, P. T. and Winter, G., *Proc. Natl. Acad. Sci. USA* 1989. *86:* 3833.
7 Hoogenboom, H. R. and Winter, G., *J. Mol. Biol.* 1992. *227:* 381.
8 Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. and Winter, G., *J. Mol. Biol.* 1992. *227:* 776.
9 Williams, S. C. and Winter, G., *Eur. J. Immunol.* 1993. *23:* 1456.
10 Huber, C., Schäble, H. F., Huber, E., Klein, R., Meindl, A., Thiebe, R., Lamm, R. and Zachau, H. G., *Eur. J. Immunol.* 1993. *23:* 2868.
11 Huber, C., Huber, E., Lautner-Rieske, A., Schäble, H. F. and Zachau, H. G., *Eur. J. Immunol.* 1993. *23:* 2860.
12 Malcolm, S., Barton, P., Murphy, C., Ferguson-Smith, M. A., Bentley, D. L. and Rabbitts, T. H., *Proc. Natl. Acad. Sci. USA* 1982. *79:* 4957.
13 McBride, O. W., Hieter, P. A., Hollis, G. F., Swan, D., Otey, M. C. and Leder, P., *J. Exp. Med.* 1982. *155:* 1480.
14 Meindl, A., Klobeck, H.-G., Ohnheiser, R. and Zachau, H. G., *Eur. J. Immunol.* 1990. *20:* 1855.
15 Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. and Foeller, C., *Sequences of Proteins of Immunological Interest,* U.S. Department of Health and Human Services, NIH, USA 1991.
16 Klobeck, H.-G., Bornkamm, G. W., Combriato, G., Mocikat, R., Pohlenz, H.-D. and Zachau, H. G., *Nucleic Acids Res.* 1985. *13:* 6515.
17 Lautner-Rieske, A., Huber, C., Meindl, A., Pargent, W., Schäble, K. F., Thiebe, R., Zocher, I. and Zachau, H. G., *Eur. J. Immunol.* 1992. *22:* 1023.
18 Zachau, H. G., *The Immunoglobulin Genes,* Academic Press, London and San Diego 1989, p. 91.
19 Bentley, D. L. and Rabbitts, T. H., *Cell* 1983. *32:* 181.
20 Pech, M., Smola, H., Pohlenz, H.-D., Straubinger, B., Gerl, R. and Zachau, H. G., *J. Mol. Biol.* 1985. *183:* 291.
21 Pargent, W., Schäble, K. F. and Zachau, H. G., *Eur. J. Immunol.* 1991. *21:* 1829.
22 Zachau, H. G., *Biol. Chem. Hoppe Seyler* 1990. *371:* 1.
23 Scott, M. G., Crimmins, D. L., McCourt, D. W., Zocher, I., Thiebe, R., Zachau, H. G. and Nahm, M. H., *J. Immunol.* 1989. *143:* 4110.
24 Straubinger, B., Huber, E., Lorenz, W., Osterholzer, E., Pargent, W., Pech, M., Pohlenz, H.-D., Zimmer, F.-J. and Zachau, H. G., *J. Mol. Biol.* 1988. *199:* 23.
25 Jaenichen, H.-R., Pech, M., Lindenmaier, W., Wildgruber, N. and Zachau, H. G., *Nucleic Acids Res.* 1984. *12:* 5249.
26 Straubinger, B., Thiebe, R., Huber, C., Osterholzer, E. and Zachau, H. G., *Biol. Chem. Hoppe Seyler* 1988. *369:* 601.
27 [illegible] Y., Schrantz, R., Liu, F. T. and Carson, D. A., *Proc. Natl. Acad. Sci. USA* 1986. *83:* 8318.
28 Radoux, V., Chen, P. P., Sorge, J. A. and Carson, D. A., *J. Exp. Med.* 1986. *164:* 2119.
29 Pargent, W., Meindl, A., Thiebe, R., Mitzel, S. and Zachau, H. G., *Eur. J. Immunol.* 1991. *21:* 1821.
30 Stavnezer, J., Kekish, O., Batter, D., Grenier, J., Balazs, I., Henderson, E. and Zegers, B. J., *Nucleic Acids Res.* 1985. *13:* 3495.
31 Bentley, D. L. and Rabbitts, T. H., *Nature* 1980. *288:* 730.
32 Pech, M. and Zachau, H. G., *Nucleic Acids Res.* 1984. *12:* 9229.
33 Scott, M. G., Crimmins, D. L., McCourt, D. W., Chung, G., Schäble, K. F., Thiebe, R., Quenzel, E.-M., Zachau, H. G. and Nahm, M. H., *J. Immunol.* 1991. *147:* 4007.
34 Liu, M. F., Robbins, D. L., Crowley, J. J., Sinha, S., Kozin, F., Kipps, T. J., Carson, D. A. and Chen, P. P., *J. Immunol.* 1989. *142:* 688.
35 Pech, M., Jaenichen, H.-R., Pohlenz, H.-D., Neumaier, P. S., Klobeck, H.-G. and Zachau, H. G., *J. Mol. Biol.* 1984. *176:* 189.
36 Lorenz, W., Schäble, K. F., Thiebe, R., Stavnezer, J. and Zachau, H. G., *Mol. Immunol.* 1988. *25:* 479.
37 Pohlenz, H.-D., Straubinger, B., Thiebe, R., Pech, M., Zimmer, F.-J. and Zachau, H. G., *J. Mol. Biol.* 1987. *193:* 241.
38 Zimmer, F.-J., Huber, C., Quenzel, M., Schek, H., Stiller, C., Thiebe, R. and Zachau, H. G., *Biol. Chem. Hoppe Seyler* 1990. *371:* 283.
39 Huber, C., Thiebe, R., Hameister, H., Smola, H., Lötscher, E. and Zachau, H. G., *Nucleic Acids Res.* 1990. *18:* 3475.
40 Lötscher, E., Grzeschik, K.-H., Bauer, H. G., Pohlenz, H.-D., Straubinger, B. and Zachau, H. G., *Nature* 1986. *320:* 456.
41 Lötscher, E., Zimmer, F.-J., Klopstock, T., Grzeschik, K.-H., Jaenichen, R., Straubinger, B. and Zachau, H. G., *Gene* 1988. *69:* 215.
42 Straubinger, B., Thiebe, R., Pech, M. and Zachau, H. G., *Gene* 1988. *69:* 209.
43 Röschenthaler, F., Schäble, K. F., Thiebe, R. and Zachau, H. G., *Biol. Chem. Hoppe Seyler* 1992. *373:* 177.
44 Miller, J. H., *Experiments in Molecular Genetics,* Cold Spring Harbor Laboratory Press, New York 1972.
45 Sanger, F., Nicklen, S. and Coulson, A. R., *Proc. Natl. Acad. Sci. USA* 1977. *74:* 5463.
46 Zachau, H. G., *Abstracts of the 2nd European Human Genome Organisation Conference,* Porto Conte, Italy, April 1992, HUGO Europe, Italy 1992, p. 11.
47 Maruyama, I. N., *Techniques – J. Methods Cell. Mol. Biol.* 1990. *2:* 302.
48 Shuldiner, A. R., Nirula, A. and Roth, J., *Nucleic Acids Res.* 1989. *17:* 4409.
49 Timmers, E., Hermans, M. M., Kraakman, M. E., Hendriks, R. W. and Schuurman, R. K., *Eur. J. Immunol.* 1993. *23:* 619.
50 Chothia, C. and Lesk, A. M., *J. Mol. Biol.* 1987. *196:* 901.
51 Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith, G. S., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M. and Polyak, R. J., *Nature* 1989. *342:* 877.

ARTHRITIS & RHEUMATISM Volume 36
Number 12, December 1993, pp. 1681–1690

C8

TREATMENT OF RHEUMATOID ARTHRITIS WITH CHIMERIC MONOCLONAL ANTIBODIES TO TUMOR NECROSIS FACTOR α

MICHAEL J. ELLIOTT, RAVINDER N. MAINI, MARC FELDMANN, ALICE LONG-FOX, PETER CHARLES, PETER KATSIKIS, FIONULA M. BRENNAN, JEAN WALKER, HANNY BIJL, JOHN GHRAYEB, and JAMES N. WOODY

Objective. **To evaluate the safety and efficacy of a chimeric monoclonal antibody to tumor necrosis factor α (TNFα) in the treatment of patients with rheumatoid arthritis (RA).**

Methods. **Twenty patients with active RA were treated with 20 mg/kg of anti-TNFα in an open phase I/II trial lasting 8 weeks.**

Results. **The treatment was well tolerated, with no serious adverse events. Significant improvements were seen in the Ritchie Articular Index, which fell from a median of 28 at study entry to a median of 6 by week 6 ($P < 0.001$), the swollen joint count, which fell from 18 to 5 ($P < 0.001$) over the same period, and in the other major clinical assessments. Serum C-reactive protein levels fell from a median of 39.5 mg/liter at study entry to 8 mg/liter at week 6 ($P < 0.001$), and significant decreases were also seen in serum amyloid A and interleukin-6 levels.**

Conclusion. **Treatment with anti-TNFα was safe and well tolerated and resulted in significant clinical and laboratory improvements. These preliminary results support the hypothesis that TNFα is an important regulator in RA, and suggest that it may be a useful new therapeutic target in this disease.**

Despite optimal use of current antirheumatic therapy, the outcome for many patients with rheumatoid arthritis (RA) consists of pain, disability, and premature death (1–3). As a response to the need for more effective and less toxic treatment, and to an increase in our understanding of the pathogenic mechanisms in RA, several groups have used monoclonal antibodies as therapeutic agents in this disease (4–10). Such immunotherapy has been, in most cases, targeted specifically to the T cell, a strategy based on evidence that T cells are involved in the initiation and maintenance of RA (11).

Here, we outline an alternative immunotherapeutic strategy, which involves the use of monoclonal antibodies with specificity for a cytokine, tumor necrosis factor α (TNFα). This approach is based on a body of knowledge regarding the role of cytokines in general, and of TNFα in particular, in the inflammatory process in RA. The first clearly documented study demonstrated the presence of interleukin-1 (IL-1) in RA synovial fluid (12). Subsequently, we and others have reported the presence and local synthesis in

From the Clinical Immunology and Sunley Divisions, Kennedy Institute of Rheumatology, and the Academic Department of Rheumatology, Charing Cross and Westminster Medical School, London, United Kingdom; and Centocor Inc., Malvern, Pennsylvania.

Supported in part by grants from the Arthritis and Rheumatism Council of Great Britain. Dr. Elliott's work was supported by the Michael Mason Fellowship from the Arthritis Foundation of Australia, and by the DEV Starr Fellowship from the Royal Australasian College of Physicians.

Michael J. Elliott, MB, BS, PhD, FRACP: Clinical Immunology Division, Kennedy Institute of Rheumatology; Ravinder N. Maini, MB, BChir, FRCP: Clinical Immunology Division, Kennedy Institute of Rheumatology; Marc Feldmann, MB, BS, PhD, FRCPath: Sunley Division, Kennedy Institute of Rheumatology; Alice Long-Fox, RGN: Clinical Immunology Division, Kennedy Institute of Rheumatology; Peter Charles, FIMLS: Department of Rheumatology, Charing Cross and Westminster Medical School; Peter Katsikis, MD, PhD: Sunley Division, Kennedy Institute of Rheumatology; Fionula M. Brennan, PhD: Sunley Division, Kennedy Institute of Rheumatology; Jean Walker, RGN, RM: Clinical Immunology Division, Kennedy Institute of Rheumatology; Hanny Bijl, MD: Centocor, Inc; John Ghrayeb, PhD: Centocor Inc; James N. Woody, MD, PhD: Centocor, Inc.

Address reprint requests to Michael J. Elliott, PhD, Kennedy Institute of Rheumatology, 6 Bute Gardens, London W6 7DW, United Kingdom.

Submitted for publication May 28, 1993; accepted in revised form August 10, 1993.

Table 1. Demographic features of 20 patients with refractory rheumatoid arthritis

Patient	Age/sex	Disease duration (years)	Previous DMARDs*	Concomitant therapy†
1	48/F	7	SSZ, DP, GST, AUR, MTX, AZA, HCQ	Pred. 5 mg
2	63/F	7	SSZ, GST, DP	Para. 1–2 gm
3	59/M	3	AUR, HCQ, GST, MTX, SSZ	Pred. 10 mg, Indo. 225 mg
4	56/M	10	GST, DP, AZA, SSZ	Pred. 12.5 mg, Ibu. 2 gm, Para. 1–2 gm
5	28/F	3	GST, SSZ, DP, AZA	Pred. 8 mg, Para. 1–2 gm, Code. 16 mg
6	40/M	3	SSZ, HCQ, AUR	Nap. 1 gm
7	54/F	7	DP, GST, SSZ, AZA, MTX	Para. 1–2 gm, Code. 16–32 mg
8	23/F	11	HCQ, GST, SSZ, MTX, AZA	Pred. 7.5 mg, Dicl. 100 mg, Para. 1–2 gm, Dext. 100–200 mg
9	51/F	15	GST, HCQ, DP, MTX	Pred. 7.5 mg, Dicl. 125 mg, Para. 1–3 gm
10	47/F	12	SSZ, CYC, MTX	Ben. 4 gm
11	34/F	10	DP, SSZ, MTX	Pred. 10 mg, Para. 1.5 gm, Code. 30–90 mg
12	57/F	12	GST, MTX, DP, AUR	Asp. 1.2 gm
13	51/F	7	SSZ, AZA	Para. 1–4 gm
14	72/M	11	GST, DP, AZA, MTX	Pred. 5 mg, Para. 1–4 gm, Code. 16–64 mg
15	51/F	17	HCQ, DP, SSZ, MTX	Asp. 0.3 gm
16	62/F	16	GST, DP, SSZ, MTX, AZA	Para. 1–4 gm, Code. 16–64 mg
17	56/F	11	SSZ, DP, GST, MTX, HCQ, AZA	Pred. 7.5 mg, Eto. 600 mg, Para. 1–2 gm, Dext. 100–200 mg
18	48/F	14	GST, MTX, DP, SSZ, AUR, AZA	Pred. 7.5 mg, Indo. 100 mg, Para. 1–3 gm
19	42/F	3	SSZ, MTX	Fen. 450 mg, Ben. 6 gm, Code. 30 mg
20	47/M	20	GST, DP, SSZ, AZA	Pred. 10 mg, Para. 1–3 gm

* Disease-modifying antirheumatic drugs (DMARDs) were SSZ = sulfasalazine; DP = D-penicillamine; GST = gold sodium thiomalate; AUR = auranofin; MTX = methotrexate; AZA = azathioprine; HCQ = hydroxychloroquine; CYC = cyclophosphamide.
† Daily doses are shown. Pred. = prednisolone; Para. = paracetamol; Indo. = indomethacin; Ibu. = ibuprofen; Code. = codeine phosphate; Nap. = naprosyn; Dicl. = diclofenac; Dext. = dextropropoxyphene; Ben. = benorylate; Asp. = aspirin; Eto. = etodolac; Fen. = fenbufen.

rheumatoid synovial membrane of many cytokines, including IL-1 (13), TNFα (13,14), IL-6 (15), granulocyte-macrophage colony-stimulating factor (GM-CSF; 16), IL-8 (17), and transforming growth factor β (TGFβ) (18,19).

We have investigated the relationships between these cytokines in RA, using a synovial culture system in which dissociated rheumatoid synovial cells are allowed to spontaneously re-aggregate in vivo. Even in the absence of extrinsic stimulation, such cells express high levels of cytokines and HLA class II molecules (20). Using this system, we showed that production of bioactive IL-1 was abrogated by neutralizing antibodies to TNFα, but not by antibodies to TNFβ or by normal rabbit IgG (21). This occurred in rheumatoid, but not osteoarthritic, cultures and suggested to us that TNFα was of particular importance as a regulatory cytokine. Subsequent analysis reinforced this concept, with the demonstration that another proinflammatory cytokine, GM-CSF, was regulated in the synovial membrane by TNFα (22) and that TNFα receptor expression, necessary for transmitting TNFα signals, was up-regulated in rheumatoid synovium (23,24).

Two recent mouse studies provide further insight into the importance of TNFα in arthritis. Keffer et al (25) described a mouse transgenic for the human TNFα gene, which expressed high levels of human TNFα in vivo and which reproducibly developed arthritis beginning at 4 weeks of age. The disease in these animals could be prevented by administration of monoclonal antibodies to human TNFα. In separate experiments in our own laboratory, we showed that in the type II collagen arthritis model in the DBA/1 mouse, the hamster anti-murine TNF monoclonal antibody TN3.19.2 significantly ameliorated the inflammation and tissue destruction when administered before or after the onset of disease (26).

Based on these considerations, it was of interest to determine the effect of therapy with a chimeric (human IgG1, murine Fv) monoclonal antibody to human TNFα in patients with rheumatoid arthritis. We report here that anti-TNFα therapy was safe and well tolerated, and induced marked improvements in both clinical and laboratory disease measures. These findings are consistent with our postulate concerning the critical role of TNFα in the pathogenesis of RA (27,28), and suggest that TNFα may be a useful therapeutic target in this disease.

Table 2. Changes in clinical assessments following treatment of rheumatoid arthritis patients with cA2*

Week of trial	Morning stiffness, minutes	Pain score, 0–10 cm	Ritchie index, 0–69	Swollen joint count, 0–28	Grip strength, 0–300 mm Hg		IDA, 1–4	Patient's assessment, no. grades improved, 0–3
					Left hand	Right hand		
Screen	135, 0–600	7.4, 4–9.7	23, 4–51	16, 4–28	84, 45–300	96, 57–300	3, 2.3–3.3	NA
0	180, 20–600	7.1, 2.7–9.7	28, 4–52	18, 3–27	77, 52–295	92, 50–293	3, 2–3.5	NA
1	20, 0–180 (<0.001†)	2.6, 0.6–7.8 (<0.001†)	13, 2–28 (<0.001; <0.002†)	13.5, 1–25 (>0.05)	122, 66–300 (>0.05)	133, 57–300 (>0.05)	2, 1.5–3.3 (<0.001†)	1, 1–3
2	15, 0–150 (<0.001†)	3.0, 0.3–6.4 (<0.001†)	13, 1–28 (<0.001†)	11.5, 1–22 (<0.003; <0.02†)	139, 75–300 (<0.03; >0.05†)	143, 59–300 (>0.05)	2, 1.5–3.2 (<0.001†)	1.5, 1–3
3	5, 0–150 (<0.001†)	2.2, 0.2–7.4 (<0.001†)	8, 0–22 (<0.001†)	6, 1–19 (<0.001; <0.002†)	113, 51–300 (>0.05)	142, 65–300 (>0.05)	2, 1.2–3.2 (<0.001†)	2, 1–2
4	15, 0–90 (<0.001†)	1.9, 0.1–5.6 (<0.001†)	10, 0–17 (<0.001†)	6, 0–21 (<0.001; <0.002†)	124, 79–300 (<0.02; >0.05†)	148, 64–300 (<0.03; >0.05†)	1.8, 1.3–2.7 (<0.001†)	2, 1–2
6	5, 0–90 (<0.001†)	1.9, 0.1–6.2 (<0.001†)	6, 0–18 (<0.001†)	5, 1–14 (<0.001†)	119, 68–300 (<0.04; >0.05†)	153, 62–300 (<0.05; >0.05†)	1.7, 1.3–2.8 (<0.001†)	2, 1–2
8	15, 0–60 (<0.001†)	2.1, 0.2–7.7 (<0.001†)	8, 1–28 (<0.001†)	7, 1–18 (<0.001†)	117, 69–300 (<0.03; >0.05†)	167, 53–300 (<0.03; >0.05†)	1.8, 1.5–2.8 (<0.001†)	2, 1–3

* Values are the median, range (*P*) for 20 patients for the initial screen and weeks 0–2, and for 19 patients thereafter. Patient 15 dropped out after week 2 of study. All *P* values versus week 0, by Mann-Whitney test. IDA = Index of Disease Activity; NA = not applicable.
† Adjusted for multiple statistical comparisons.

PATIENTS AND METHODS

Patient selection. Twenty patients were recruited, each of whom fulfilled the American College of Rheumatology (formerly, the American Rheumatism Association) criteria for the diagnosis of RA (29). The clinical characteristics of the patients are shown in Table 1. The study group comprised 15 females and 5 males, with a median age of 51 years (range 23–72), a median disease duration of 10.5 years (range 3–20), and a history of failed therapy with standard disease-modifying antirheumatic drugs (DMARDs) (median number of failed DMARDs 4, range 2–7).

Seventeen patients were seropositive at study entry or had been seropositive at some stage of their disease. All had erosions evident on radiographs of the hands or feet, and 3 had rheumatoid nodules. All patients had active disease at trial entry, as defined by an Index of Disease Activity (IDA) (30) of at least 1.75, together with at least 3 swollen joints, and were classified in anatomic and functional stage II or III (31). The pooled data for each of the clinical and laboratory indices of disease activity at the time of screening for the trial (up to 4 weeks prior to trial entry), and on the day of trial entry itself (week 0), are shown in Tables 2 and 3.

All DMARDs were discontinued at least 1 month prior to trial entry. Patients were allowed to continue taking a nonsteroidal antiinflammatory drug and/or prednisolone (≤12.5 mg/day) during the trial. The dosage of these agents was kept stable for 1 month prior to trial entry and during the course of the trial. No parenteral corticosteroids were allowed during these periods. Simple analgesics were allowed ad libitum.

Patients with other serious medical conditions were excluded from study. Specific exclusions were as follows: serum creatinine >150 μmoles/liter (normal 60–120), hemoglobin (Hgb) <90 gm/liter (normal 120–160 in females, and 135–175 in males), white blood cell (WBC) count <4 × 10^9/liter (normal 4–11 × 10^9/liter), platelet count <100 × 10^9/liter (normal 150–400 × 10^9/liter), and abnormal liver enzyme levels or active pathology noted on chest radiographs.

All patients gave their informed consent for the trial, and approval was granted by the local ethics committee.

Treatment protocol. cA2 is a chimeric human/mouse monoclonal anti-TNFα antibody, consisting of the constant regions of human (Hu)IgG1κ, coupled to the Fv region of a high-affinity neutralizing murine anti-HuTNFα antibody (A2). The antibody was produced by Centocor Inc., by continuous fermentation of a mouse myeloma cell line which had been transfected with cloned DNA coding for cA2, and was purified from culture supernatant by a series of steps involving column chromatography. The chimeric antibody retains specificity for natural and recombinant HuTNFα, and is of high affinity.

The antibody was stored at 4°C in 20-ml vials containing 5 mg of cA2 per milliliter of 0.01*M* phosphate buffered saline in 0.15*M* sodium chloride at a pH of 7.2 and was filtered through a 0.2-μm sterile filter before use. The appropriate amount of cA2 was then diluted to a total volume of 300 ml in sterile saline and administered intravenously via a 0.2-μm in-line filter over a period of 2 hours.

Patients were admitted to the hospital for 8–24 hours for each treatment, and were mobile except during infusions. The trial was of an open, uncontrolled design, with a comparison of 2 treatment schedules. Patients 1–5 and 11–20 received a total of 2 infusions, each consisting of 10 mg/kg of cA2, at entry to the study (week 0) and 14 days later (week 2). Patients 6–10 received a total of 4 infusions of 5 mg/kg at cA2, at entry and on days 4, 8, and 12. The total dose received by the 2 patient groups was therefore the same: 20 mg/kg.

Assessments. *Safety monitoring.* Vital signs were recorded every 15–30 minutes during infusions, and at intervals for up to 24 hours postinfusion. Patients were

Table 3. Changes in laboratory measures following treatment of rheumatoid arthritis patients with cA2*

Week of trial	Hgb, gm/liter	WBC, $\times 10^9$/liter	Platelets, $\times 10^9$/liter	ESR, mm/hour	CRP, mg/liter	SAA, mg/ml	RF, inverse titer
Screen	117, 98–146	7.9, 3.9–15.2	352, 274–631	59, 18–87	42, 9–107	ND	ND
0	113, 97–144	9.0, 4.9–15.7	341, 228–710	55, 15–94	39.5, 5–107	245, 18–1,900	2,560, 160–10,240
1	114, 96–145 (>0.05)	8.5, 3.6–13.6 (>0.05)	351, 223–589 (>0.05)	26, 13–100 (>0.05)	5, 0–50 (<0.001†)	58, 0–330 (<0.001; <0.003†)	ND
2	112, 95–144 (>0.05)	8.2, 4.3–12.7 (>0.05)	296, 158–535 (<0.04; >0.05†)	27, 10–90 (<0.02; >0.05†)	5.5, 0–80 (<0.001; <0.003†)	80, 11–900 (<0.02; <0.04†)	ND
3	110, 89–151 (>0.05)	9.0, 3.7–14.4 (>0.05)	289, 190–546 (<0.03; >0.05†)	27, 12–86 (<0.04; >0.05†)	7, 0–78 (<0.001; <0.002†)	ND	ND
4	112, 91–148 (>0.05)	8.2, 4.7–13.9 (>0.05)	314, 186–565 (>0.05)	23, 10–87 (<0.04; >0.05†)	10, 0–91 (<0.004; <0.02†)	ND	ND
6	116, 91–159 (>0.05)	9.1, 2.9–13.9 (>0.05)	339, 207–589 (>0.05)	23, 12–78 (<0.03; >0.05†)	8, 0–59 (<0.001†)	ND	ND
8	114, 94–153 (>0.05)	7.6, 4.2–13.5 (>0.05)	339, 210–591 (>0.05)	30, 7–73 (>0.05)	6, 0–65 (<0.001†)	ND	480, 40–5,120 (>0.05)

* Values are the median, range (*P*) for 20 patients for the initial screen and weeks 0–2, and for 19 patients thereafter. Patient 15 dropped out after week 2 of study. For rheumatoid factor (RF), only those patients with week 0 titers ≥1:160 in the particle agglutination assay were included (n = 14). All *P* values versus week 0, by Mann-Whitney test. Normal ranges: hemoglobin (Hgb) 120–160 gm/liter in females and 135–175 gm/liter in males; white blood cell (WBC) count 4–11 × 10^9/liter; platelet count 150–400 × 10^9/liter; erythrocyte sedimentation rate (ESR) <15 mm/hour in females and <10 mm/hour in males; C-reactive protein (CRP) <10 mg/liter; serum amyloid A (SAA) <10 mg/ml. ND = not done.

questioned concerning possible adverse events before each infusion and at weeks 1, 2, 3, 4, 6, and 8 of the trial. A complete physical examination was performed at screening and at week 8. In addition, patients were monitored by standard laboratory tests including a complete blood cell count, and levels of C3 and C4 components of complement, IgG, IgM, and IgA, serum electrolytes, creatinine, urea, alkaline phosphatase, aspartate transaminase, and total bilirubin.

Sample times for these tests were between 0800 and 0900 hours (preinfusion) and 1200–1400 hours (24 hours postinfusion). Blood tests subsequent to day 1 were performed in the morning, usually between 0700 and 1200 hours. Urine analysis and culture were also performed at each assessment point.

Response assessment. The patients were assessed for response to cA2 at weeks 1, 2, 3, 4, 6, and 8 of the trial. The assessments were all made between 0700 and 1300 hours by the same observer (AL-F). The following clinical assessments were made: duration of morning stiffness (minutes), pain score (0–10 cm on a visual analog scale), Ritchie Articular Index (maximum score 69) (32), number of swollen joints (28 joint count) (validation described in ref. 33), grip strength (0–300 mm Hg, mean of 3 measurements per hand, by sphygmomanometer cuff), and an assessment of function (the Stanford Health Assessment Questionnaire [HAQ], modified for British patients [34]). In addition, the patients' global assessments of response were recorded using a 5-point scale (worse, no response, fair response, good response, excellent response).

Routine laboratory indicators of disease activity included complete blood cell counts, C-reactive protein (CRP) levels (by rate nephelometry), and the erythrocyte sedimentation rate (ESR; Westergren). Followup assessments were made at monthly intervals after the conclusion of the formal trial period, in order to assess the duration of response.

Analysis of improvement in individual patients was made using two separate indices. First, an IDA was calculated for each time point according to the method of Mallya and Mace (30), with input variables of morning stiffness, pain score, Ritchie Articular Index, grip strength, ESR, and Hgb. The second index calculated was that of Paulus et al (35), which uses input variables of morning stiffness, ESR, joint pain/tenderness, joint swelling, and patient's and physician's global assessments of disease severity.

To calculate the presence (or otherwise) of a response according to this index, two approximations were made to accommodate our data. The swollen joint count used by us (nongraded total of swollen joints of 28 joints assessed), which has been validated (33), was used in place of the more extensive graded count described by Paulus et al, and the patient's and physician's global assessments of *response* recorded by us were approximated to the global assessments of *disease activity* used by Paulus et al (35). In addition to determining response according to these published indices, we selected 6 disease activity assessments of interest (morning stiffness, pain score, Ritchie Articular Index, swollen joint count, ESR, and CRP) and calculated their mean percentage improvement. We have used this value to give an indication of the degree of improvement seen in responding patients.

Immunologic investigations. Rheumatoid factors were measured using the rheumatoid arthritis particle agglutination assay (RAPA) (FujiBerio Inc, Tokyo, Japan), in which titers of 1:160 or greater were considered significant. Rheumatoid factor isotypes were measured by enzyme-linked immunosorbent assay (ELISA) (Cambridge Life Sciences, Ely, UK). Addition of cA2, at concentrations of up to 200 μg/ml, to these assay systems did not alter the assay results (data not shown).

Antinuclear antibodies were detected by immunoflu-

orescence on HEp-2 cells (Biodiagnostics, Upton, UK), and antibodies to extractable nuclear antigens were measured by counterimmunoelectrophoresis with polyantigen extract (Biodiagnostics). Sera positive by immunofluorescence were also screened for antibodies to DNA by the Farr assay (Kodak Diagnostics, Amersham, UK). Anticardiolipin antibodies were measured by ELISA (Shield Diagnostics, Dundee, Scotland). Serum amyloid A (SAA) was measured by sandwich ELISA (Biosource International, Camarillo, CA). Antiglobulin responses to the infused chimeric antibody were measured by an in-house ELISA, using cA2 as a capture reagent.

Cytokine assays. Bioactive TNF was measured in sera using the WEHI 164 clone 13 cytotoxicity assay (36). Total IL-6 was measured in sera using a commercial immunoassay (Medgenix Diagnostics, Brussels, Belgium) and using a sandwich ELISA developed in-house, with monoclonal antibodies provided by Dr. F. di Padova (Basel, Switzerland). Microtiter plates were coated with monoclonal antibody LNI 314-14 at a concentration of 3 μg/ml for 18 hours at 4°C, and blocked with 3% bovine serum albumin in 0.1*M* phosphate buffered saline, pH 7.2. Undiluted sera or standards (recombinant HuIL-6, 0–8.1 μg/ml) were added to the wells in duplicate and incubated for 18 hours at 4°C. Bound IL-6 was detected by incubation with monoclonal antibody LNI 110-14 for 90 minutes at 37°C, followed by biotin-labeled goat anti-murine IgG2b for 90 minutes at 37°C (Southern Biotechnology, Birmingham, AL). The assay was developed using streptavidin–alkaline phosphatase (Southern Biotechnology) and *p*-nitrophenyl phosphate as a substrate, and the optical density read at 405 nm.

Statistical analysis. Data for week 0 versus subsequent time points were compared for each assessment using the Mann-Whitney test. For comparison of rheumatoid factor titers (by RAPA), the data were expressed as dilutions before applying the test.

This was an exploratory study, in which prejudgments about the optimal times for assessment were not possible. Although it has not been common practice to adjust for multiple statistical comparisons in such studies (4–10), a conservative statistical approach would require adjustment of *P* values to take into account analysis at several time points. The *P* values have therefore been presented in two forms: unadjusted, and after making allowance for analysis at multiple time points by use of the Bonferroni adjustment. Where *P* values remained <0.001 after adjustment, a single value only is given. A *P* value of <0.05 is considered significant.

RESULTS

Safety of cA2. The administration of cA2 was exceptionally well tolerated, with no headache, fever, hemodynamic disturbance, allergy, or other acute manifestation. No serious adverse events were recorded during the 8-week trial. Two minor infective episodes were recorded, each "possibly related" to cA2: patient 15 presented at week 2 with clinical features of bronchitis. Sputum culture grew only normal commensals. She had a history of smoking and of a similar illness 3 years previously. The illness responded promptly to treatment with amoxicillin, but her second cA2 infusion was withheld and the data for this patient are therefore not analyzed beyond week 2. Patient 18 showed significant bacteriuria on routine culture at week 6 (>10^5/ml; lactose-fermenting coliform), but was asymptomatic. This condition also responded promptly to amoxicillin.

Routine analysis of blood samples showed no consistent adverse changes in hematologic parameters, renal function, liver function, or levels of C3, C4, or immunoglobulins during the 8 weeks of the trial. Four minor, isolated, and potentially adverse laboratory disturbances were recorded. Patient 2 experienced a transient rise in blood urea levels, from 5.7 mmoles/liter to 9.2 mmoles/liter (normal 2.5–7), with no change in serum creatinine. This change was associated with the temporary use of a diuretic, which had been prescribed for a non-rheumatologic disorder. The value normalized within 1 week and was classified as "probably not related" to cA2.

Patient 6 experienced a transient fall in the peripheral blood lymphocyte count, from 1.6 × 10^9/liter to 0.8 × 10^9/liter (normal 1.0–4.8). This abnormality was not seen at the next sample point (2 weeks later), was not associated with any clinical manifestations, and was classified as "possibly related" to cA2. Patients 10 and 18 developed elevated titers of anti-DNA antibodies at weeks 6 and 8 of the trial. Elevated anticardiolipin antibodies were also detected in patient 10. Both patients had a preexisting positive antinuclear antibody titer, and patient 10 had a history of borderline lymphocytopenia and high serum IgM. There were no clinical features of systemic lupus erythematosus, and the laboratory changes were judged "probably related" to cA2.

Efficacy of cA2. The pattern of response for each of the clinical assessments of disease activity and the derived IDA are shown in Table 2. All clinical assessments showed improvement following treatment with cA2, with maximal responses from week 3. Duration of morning stiffness decreased from a median of 180 minutes at study entry (week 0) to 5 minutes at week 6 ($P < 0.001$ by Mann-Whitney test, adjusted), representing a 97% improvement. The pain score decreased from 7.1 to 1.9 over the same period ($P < 0.001$, adjusted), representing an improvement of 73%. Similarly, the Ritchie Articular Index improved from 28 to 6 at week 6 ($P < 0.001$, adjusted; 79% improvement), and the swollen joint count decreased from 18

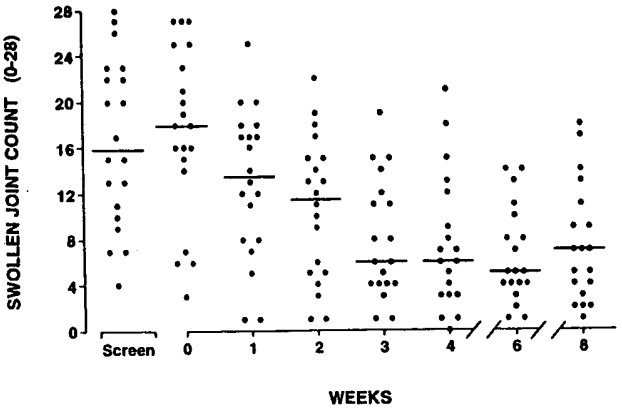


Figure 1. Swollen joint counts (maximum 28), as recorded by a single observer, in 20 patients with rheumatoid arthritis treated with cA2. The screening time point was within 4 weeks of entry to the study (week 0); data from patient 15 were not included after week 2 (dropout). Significance of the changes, relative to week 0, were determined by Mann-Whitney test (adjusted): $P > 0.05$ at week 1, $P < 0.02$ at week 2, $P < 0.002$ at weeks 3 and 4, and $P < 0.001$ at weeks 6 and 8. Bars show median values.

to 5 ($P < 0.001$, adjusted; 72% improvement). The individual swollen joint counts for all time points are shown in Figure 1.

Grip strength also improved; the median grip strength rose from 77 mm Hg (left) and 92 mm Hg (right) at week 0 to 119 (left) and 153 (right) at week 6 ($P < 0.04$ and $P < 0.05$, left and right hands, respectively; $P > 0.05$ both hands, adjusted for multiple comparisons). The IDA has a range of 1 (normal) to 4 (severe disease activity). The IDA showed a decrease from a median of 3 at study entry to 1.7 at week 6 ($P < 0.001$, adjusted). Patients were asked to grade their responses to cA2 using a 5-point scale. No patient recorded a response of "worse" or "no change" at any point in the trial. "Fair," "good," and "excellent" responses were classified as improvements of 1, 2, and 3 grades, respectively. At week 6, there was a median of 2 grades of improvement (Table 2).

We also measured changes in the patients' functional capacity, using the HAQ, as modified for British patients (range 0–3). The median (range) HAQ score improved from 2 (0.9–3) at study entry to 1.1 (0–2.6) by week 6 ($P < 0.001$ and $P < 0.002$ adjusted).

The changes in the laboratory values which reflect disease activity are shown in Table 3. The most rapid and impressive changes were seen in serum CRP levels, which fell from a median of 39.5 mg/liter at week 0 (normal <10) to 8 mg/liter by week 6 of the trial ($P < 0.001$, adjusted), representing an improvement of 80%. Of the 19 patients with elevated CRP at study entry, 17 showed decreases to the normal range at some point during the trial. The improvement in CRP was maintained in most patients over the assessment period (Table 3 and Figure 2); the exceptions with high values at 4 and 6 weeks tended to be those with the highest starting values (data not shown).

The ESR also showed improvement, with a fall from 55 mm/hour at study entry (normal <10 in males and <15 in females) to 23 mm/hour at week 6 ($P < 0.03$ and $P > 0.05$ adjusted; 58% improvement). SAA levels were elevated in all patients at trial entry, and fell from a median of 245 mg/ml (normal <10) to 58 mg/ml at week 1 ($P < 0.003$ adjusted; 76% improvement) and to 80 mg/ml at week 2 ($P < 0.04$, adjusted). No significant changes were seen in Hgb level, WBC count, or platelet count at week 6, although the platelet count did improve at weeks 2 and 3 compared with trial entry (Table 3).

The response data were also analyzed for each patient individually (not shown). The majority of patients had their best overall responses at week 6, at which time 13 assessed their responses as "good" while 6 assessed their responses as "fair." Eighteen of the 19 patients who completed the treatment schedule achieved an improvement in the IDA of 0.5 or greater at week 6, and 10 achieved an improvement of 1.0 or greater. All patients achieved a response at week 6

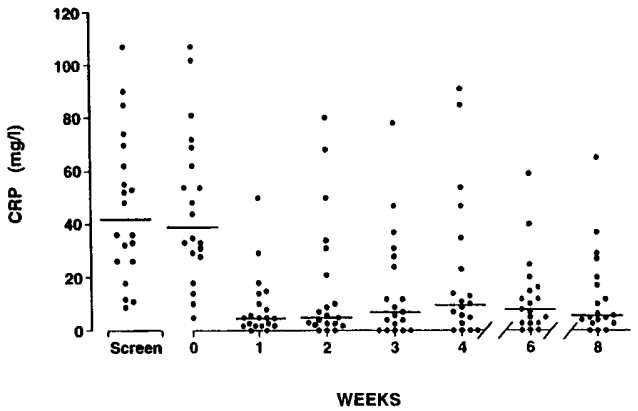


Figure 2. Serum C-reactive protein (CRP) levels (normal 0–10 mg/liter), as measured by nephelometry, in 20 patients with rheumatoid arthritis treated with cA2. The screening time point was within 4 weeks of entry to the study (week 0); data from patient 15 were not included after week 2 (dropout). Significance of the changes, relative to week 0, were determined by Mann-Whitney test (adjusted): $P < 0.001$ at week 1, $P < 0.003$ at week 2, $P < 0.002$ at week 3, $P < 0.02$ at week 4, and $P < 0.001$ at weeks 6 and 8. Bars show median values.

according to the index described by Paulus et al (35). At week 6, all patients showed a mean improvement of 30% or greater in the 6 selected measures of disease activity (see Patients and Methods), with 18 of the 19 patients showing a mean improvement of 50% or greater (data not shown).

Although the study was primarily designed to assess the short-term effects of cA2 treatment, followup clinical and laboratory data are available for those patients followed for sufficient time (n = 12). The duration of response in these patients, defined as the duration of a 30% (or greater) mean improvement in the 6 selected disease activity measures, was variable, ranging from 8 weeks to 25 weeks (median 14) (data not shown).

Comparison of the clinical and laboratory data for patients treated with 2 infusions of cA2 (each at 10 mg/kg) versus those treated with 4 infusions (each at 5 mg/kg) showed no significant differences in the rapidity or extent of response (data not shown).

Immunologic investigations and cytokines. Measurement of rheumatoid factor by RAPA showed 14 patients with significant titers (≥1:160) at trial entry. Of these, 6 patients showed a decrease of at least 2 titers on treatment with cA2, while the remaining patients showed a change of 1 titer or less. No patient showed a significant increase in rheumatoid factor titer during the trial (data not shown). The median titer in the 11 patients decreased from 1:2,560 at entry to 1:480 by week 8 ($P > 0.05$) (Table 3). Specific rheumatoid factor isotypes were measured by ELISA, and showed decreases in the 6 patients whose RAPA had declined significantly, as well as in some other patients (data not shown). Median values for the 3 isotypes in the 14 patients seropositive at trial entry were 119, 102, and 62 IU/ml (IgM, IgG, and IgA isotypes, respectively) and at week 8 were 81, 64, and 46 IU/ml ($P > 0.05$).

We tested sera from patients 1–9 for the presence of bioactive TNF, using the WEHI 164 clone 13 cytotoxicity assay (36). In 8 patients, serum samples spanning the entire trial period were tested, while for patient 9, only 3 samples (1 pretrial, 1 intermediate, and the last available sample) were tested. The levels of bioactive TNF were below the limit of sensitivity of the assay in the presence of human serum (1 pg/ml) (data not shown).

Since production of CRP and SAA are thought to be regulated in large part by IL-6, we also measured serum levels of this cytokine, using 2 different assays which measure total IL-6. In the Medgenix assay, IL-6 was significantly elevated in 17 of the 20 patients at study entry. In this group, levels fell from 60 pg/ml (range 18–500) to 40 pg/ml (range 0–230) at week 1 ($P > 0.05$) and to 32 pg/ml (range 0–210) at week 2 ($P < 0.005$ and $P < 0.01$, adjusted). These results were supported by measurement of serum IL-6 in the first 16 patients in a separate ELISA developed in-house. IL-6 was detectable in 11 of these samples, with median (range) levels falling from 210 pg/ml (25–900) at entry to 32 pg/ml (0–1,700) at week 1 ($P < 0.02$ and $P < 0.04$, adjusted) and to 44 pg/ml (0–240) at week 2 ($P < 0.02$ and $P < 0.03$, adjusted).

We tested sera from patients 1–10 for the presence of antiglobulin responses to the infused chimeric antibody, but none were detected (data not shown). In many patients, however, cA2 was still detectable in serum samples taken at week 8 (data not shown) and this may have interfered with the ELISA.

DISCUSSION

This is the first report describing the administration of anti-TNFα antibodies for treatment of human autoimmune disease. Many cytokines are produced in rheumatoid synovium, but we chose to specifically target TNFα because of mounting evidence that it was a major molecular regulator in RA (21,22,26–28). The study results presented here support that view and allow 3 important conclusions to be drawn.

First, treatment with cA2 was safe and the infusion procedure was well tolerated. Although fever, headache, chills, and hemodynamic disturbance have all been reported following treatment with anti-CD4 or anti-CDw52 in RA (6,10), such features were absent in our patients. Also notable was the absence of any allergic event despite repeated treatment with the chimeric antibody, although the interval between initial and repeat infusions may have been too short to allow maximal expression of any antiglobulin response. The continuing presence of circulating cA2 at the conclusion of the trial may have precluded detection of antiglobulin responses, but also indicated that any such responses were likely to be of low titer and/or affinity. Although we recorded 2 episodes of infection among the study group, these were minor and the clinical courses were unremarkable. TNFα has been implicated in the control of *Listeria* and other infections in mice (37), but our limited experience does not suggest an increased risk of infection after TNFα blockade in humans.

The second conclusion concerns the clinical

efficacy of cA2. The patients we treated had longstanding, erosive, and for the most part, seropositive disease, and therapy with several standard DMARDs had failed. Despite this, the major clinical assessments of disease activity and outcome (morning stiffness, pain score, Ritchie articular index, swollen joint count, and HAQ score) showed statistically significant improvement, even after adjustment for multiple comparisons. All patients graded their response as at least "fair," with the majority grading it as "good." In addition, all achieved a response according to the criteria of Paulus et al and showed a mean improvement of at least 30% in 6 selected disease activity measures. The design of the trial does not allow these results to be attributed to the action of cA2 alone. However, the extent of the clinical improvements, their consistency throughout the study group, and the parallel changes in laboratory indices of disease activity (see below) are encouraging.

The improvements in clinical assessments following treatment with cA2 appear to be at least as good as those reported following treatment of similar patients with antileukocyte antibodies (6,10), although firm conclusions concerning each of these agents will require controlled, blinded studies. The two therapeutic approaches may already be distinguished, however, by their effects on the acute-phase response, since in several studies of antileukocyte antibodies, no consistent improvements in CRP or ESR were seen (4–6,8,10). In contrast, treatment with cA2 resulted in significant decreases in serum CRP and SAA values, with normalization of values in many patients. The changes were rapid and marked, and in the case of CRP, sustained for the duration of the study (Table 3). The decreases in ESR were less marked, achieving statistical significance only when unadjusted for the number of comparisons (Table 3).

These results are consistent with current concepts that implicate TNFα in the regulation of hepatic acute-phase protein synthesis, either directly, or by control of other, secondary, cytokines such as IL-6 (38,39). To investigate the mechanism of control of the acute-phase response in our patients, we measured serum TNFα and IL-6 before and after cA2 treatment. Bioactive TNFα was not detectable in sera obtained at baseline or subsequently. In view of previous reports of variability between different immunoassays in the measurement of cytokines in biologic fluids (40), we used 2 different assays for IL-6, and both demonstrated significant decreases in serum IL-6 levels by week 2. These findings support the other objective laboratory changes induced by cA2, and provide in vivo evidence that TNFα may be a regulatory cytokine for IL-6 in this disease. Among the other laboratory tests performed, levels of rheumatoid factors fell significantly in 6 patients.

The mechanism of action of cA2 leading to the clinical responses outlined above was not established in this study. Neutralization of TNFα may have a number of beneficial consequences, including a reduction in the local release of cytokines such as IL-6 and other inflammatory mediators, and modulation of synovial endothelial/leukocyte interactions. cA2 may also bind directly to synovial inflammatory cells expressing membrane TNFα, with subsequent in situ cell lysis. Further studies should establish which actions of cA2 may be clinically important.

The results obtained in this small series have important implications, both scientifically and clinically. At the scientific level, the ability of the neutralizing antibody, cA2, to reduce acute-phase protein synthesis, reduce the production of other cytokines such as IL-6, and significantly improve the clinical state demonstrates that it is possible to interfere with the cytokine network in a useful manner without untoward effects. Due to the many functions and overlapping effects of cytokines such as IL-1 and TNFα, and the fact that cytokines induce the production of other cytokines and of themselves, there had been some pessimism as to whether targeting a single cytokine in vivo would have any beneficial effect (41,42). This view is clearly refuted. On the clinical side, the results of short-term treatment with cA2 are encouraging, and suggest that TNFα may be a useful new therapeutic target in RA.

ACKNOWLEDGMENTS

We thank Drs. R. McCloskey, M. Sanders, U. Nassander, C. Wortel, D. Zelinger, and R. Lorijn of Centocor Inc. for their assistance with trial management, Dr. F. di Padova for the kind gift of anti–IL-6 monoclonal antibodies, Drs. J. Kirwan, H. Bird, and P. Schur for reviewing the manuscript, and the following physicians for the referral of patients included in this study: S. Allard, M. Corbett, C. Erhardt, A. Keat, C. Higgins, C. Mackworth-Young, G. Room, A. So, and P. Venables. We also thank Ms A. Hogg, Ms S. Tyler, and Ms G. Harris for excellent technical assistance and Dr. C. Plater-Zyberk for help in videotaping patient interviews.

REFERENCES

1. Pincus T, Callahan LF: Taking mortality in rheumatoid arthritis seriously: predictive markers, socioeconomic status and comorbidity. J Rheumatol 13:841–845, 1986

2. Scott DL, Symmons DPM, Coulton BL, Popert AJ: Long-term outcome of treating rheumatoid arthritis: results after 20 years. Lancet 1:1108–1111, 1987
3. Erhardt CC, Mumford PA, Venables PJW, Maini RN: Factors predicting a poor life prognosis in rheumatoid arthritis: an eight year prospective study. Ann Rheum Dis 48:7–13, 1989
4. Herzog C, Walker C, Muller W, Rieber P, Reiter C, Riethmuller G, Wassmer P, Stockinger H, Madic O, Pichler W: Anti-CD4 antibody treatment of patients with rheumatoid arthritis: I. Effect on clinical course and circulating T cells. J Autoimmun 2:627–642, 1989
5. Goldberg D, Morel P, Chatenoud L, Boitard C, Menkes CJ, Bertoye P, Revillard J, Bach J: Immunological effects of high dose administration of anti-CD4 antibody in rheumatoid arthritis patients. J Autoimmun 4:617–630, 1991
6. Horneff G, Burmester GR, Emmrich F, Kalden JR: Treatment of rheumatoid arthritis with an anti-CD4 monoclonal antibody. Arthritis Rheum 34:129–140, 1991
7. Kirkham BW, Pitzalis C, Kingsley GH, Chikanza IC, Sabharwal S, Barbatis C, Grahame R, Gibson T, Amlot PL, Panayi GS: Monoclonal antibody treatment in rheumatoid arthritis: the clinical and immunological effects of a CD7 monoclonal antibody. Br J Rheumatol 30:459–463, 1991
8. Reiter C, Kakavand B, Rieber EP, Schattenkirchner M, Riethmüller G, Krüger K: Treatment of rheumatoid arthritis with monoclonal CD4 antibody M-T151: clinical results and immunopharmacologic effects in an open study, including repeated administration. Arthritis Rheum 34:525–536, 1991
9. Wendling D, Wijdenes J, Racadot E, Morel-Fourrier B: Therapeutic use of monoclonal anti-CD4 antibody in rheumatoid arthritis. J Rheumatol 18:325–327, 1991
10. Isaacs JD, Watts RA, Hazleman BL, Hale G, Keogen MT, Cobbold SP, Waldmann H: Humanised monoclonal antibody therapy for rheumatoid arthritis. Lancet 340:748–752, 1992
11. Ziff M: Rheumatoid arthritis: its present and future. J Rheumatol 17:127–133, 1990
12. Fontana A, Hengartner H, Weber E, Fehr K, Grob PJ, Cohen G: Interleukin-1 activity in the synovial fluid of patients with rheumatoid arthritis. Rheumatol Int 2:49–53, 1982
13. Buchan G, Barrett K, Turner M, Chantry D, Maini RN, Feldmann M: Interleukin-1 and tumour necrosis factor mRNA expression in rheumatoid arthritis: prolonged production of IL-1α. Clin Exp Immunol 73:449–453, 1988
14. Chu CQ, Field M, Feldmann M, Maini RN: Localization of tumor necrosis factor α in synovial tissues and at the cartilage–pannus junction in patients with rheumatoid arthritis. Arthritis Rheum 34:1125–1132, 1991
15. Hirano T, Matsuda T, Turner M, Miyasaka N, Buchan G, Tang B, Sato K, Shimizu M, Maini R, Feldmann M, Kishimoto T: Excessive production of interleukin 6/B cell stimulatory factor-2 in rheumatoid arthritis. Eur J Immunol 18:1797–1801, 1988
16. Xu WD, Firestein GS, Taetle R, Kaushansky K, Zvaifler NJ: Cytokines in chronic inflammatory arthritis, II: Granulocyte–macrophage colony-stimulating factor in rheumatoid synovial effusions. J Clin Invest 83:876–882, 1989
17. Brennan FM, Zachariae COC, Chantry D, Larsen CG, Turner M, Maini RN, Matsushima K, Feldmann M: Detection of interleukin-8 biologic activity in synovial fluids from patients with rheumatoid arthritis and production of interleukin-8 mRNA by isolated synovial cells. Eur J Rheumatol 20:2141–2144, 1990
18. Brennan FM, Chantry D, Turner M, Foxwell B, Maini RN, Feldmann M: Detection of transforming growth factor-β in rheumatoid arthritis synovial tissue: lack of effect on spontaneous cytokine production in joint cell cultures. Clin Exp Immunol 81:278–285, 1990
19. Chu CQ, Field M, Abney E, Zheng RQH, Allard S, Feldmann M, Maini RN: Transforming growth factor-β1 in rheumatoid synovial membrane and cartilage/pannus junction. Clin Exp Immunol 86:380–386, 1991
20. Kissonerghis AM, Maini RN, Feldmann M: High rate of HLA class II mRNA synthesis in rheumatoid arthritis joints and its persistence in culture: down regulation by recombinant interleukin-2. Scand J Immunol 29:73–82, 1989
21. Brennan FM, Chantry D, Jackson A, Maini RN, Feldmann M: Inhibitory effects of TNFα antibodies on synovial cell interleukin-1 production in rheumatoid arthritis. Lancet 2:244–247, 1989
22. Haworth C, Brennan FM, Chantry D, Turner M, Maini RN, Feldmann M: Expression of granulocyte–macrophage colony-stimulating factor (GM-CSF) in rheumatoid arthritis: regulation by tumour necrosis factor-α. Eur J Immunol 21:2575–2579, 1991
23. Brennan FM, Gibbons DL, Mitchell T, Cope AP, Maini RN, Feldmann M: Enhanced expression of tumour necrosis factor receptor mRNA and protein in mononuclear cells isolated from rheumatoid arthritis synovial joints. Eur J Immunol 22:1907–1912, 1992
24. Deleuran BW, Chu C-Q, Field M, Brennan FM, Mitchell T, Feldmann M, Maini RN: Localization of tumor necrosis factor receptors in the synovial tissue and cartilage–pannus junction in patients with rheumatoid arthritis: implications for local actions of tumor necrosis factor α. Arthritis Rheum 35:1170–1178, 1992
25. Keffer J, Probert L, Cazlaris H, Georgopoulos S, Kaslaris E, Kioussis D, Kollias G: Transgenic mice expressing human tumour necrosis factor: a predictive genetic model of arthritis. EMBO J 10:4025–4031, 1991
26. Williams RO, Feldmann M, Maini RN: Anti-tumour necrosis factor ameliorates joint disease in murine collagen-induced arthritis. Proc Natl Acad Sci U S A 89:9784–9788, 1992
27. Feldmann M, Brennan FM, Chantry D, Haworth C, Turner M, Abney E, Buchan G, Barrett K, Barkley D, Chu A, Field M, Maini RN: Cytokine production in the rheumatoid joint: implications for treatment. Ann Rheum Dis 49:480–486, 1990
28. Brennan FM, Maini RN, Feldmann M: TNFα: a pivotal role in rheumatoid arthritis? Br J Rheumatol 31:293–298, 1992
29. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS, Medsger TA Jr, Mitchell DM, Neustadt DH, Pinals RS, Schaller JG, Sharp JT, Wilder RL, Hunder GG: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31:315–324, 1988
30. Mallya RK, Mace BEW: The assessment of disease activity in rheumatoid arthritis using a multivariate analysis. Rheumatol Rehabil 20:14–17, 1981
31. Steinbrocker O, Traeger CH, Batterman RC: Therapeutic criteria in rheumatoid arthritis. JAMA 140:659–662, 1949
32. Ritchie DM, Boyle JA, McInnes JM, Jasani MK, Dalakos TG, Grieveson P, Buchanan WW: Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. Q J Med 147:393–406, 1968
33. Fuchs HA, Brooks RH, Callahan LF, Pincus T: A simplified twenty-eight–joint quantitative articular index in rheumatoid arthritis. Arthritis Rheum 32:531–537, 1989
34. Kirwan JR, Reeback JS: Stanford health assessment questionnaire modified to assess disability in British patients with rheumatoid arthritis. Br J Rheumatol 25:206–209, 1986
35. Paulus HE, Egger MJ, Ward JR, Williams HJ, and the Cooperative Systematic Studies of Rheumatic Diseases Group: Analysis of improvement in individual rheumatoid arthritis patients treated with disease-modifying antirheumatic drugs, based on the findings in patients treated with placebo. Arthritis Rheum 33:477–484, 1990
36. Espevik T, Nissen-Meyer J: A highly sensitive cell line, WEHI 164 clone 13, for measuring cytotoxic factor/tumour necrosis

factor from human monocytes. J Immunol Methods 95:99–105, 1986
37. Havell EA: Evidence that tumour necrosis factor has an important role in antibacterial resistance. J Immunol 143:2894–2899, 1989
38. Fong Y, Tracey KJ, Moldawer LL, Hesse DG, Manogue KB, Kenney JS, Lee AT, Kuo GC, Allison AC, Lowry SF, Cerami A: Antibodies to cachectin/tumour necrosis factor reduce interleukin-1β and interleukin-6 appearance during lethal bacteremia. J Exp Med 170:1627–1633, 1989
39. Guerne P-A, Zuraw BL, Vaughan JH, Carson DA, Lotz M: Synovium as a source of interleukin-6 in vivo: contribution to local and systemic manifestations of arthritis. J Clin Invest 83:585–592, 1989
40. Roux-Lombard P, Steiner G, and the Cytokine Consensus Study Group Of The European Workshop For Rheumatology Research: Preliminary report on cytokine determination in human synovial fluids: a consensus study of the European Workshop for Rheumatology Research. Clin Exp Rheumatol 10:515–520, 1992
41. Kingsley G, Panayi G, Lanchbury J: Immunotherapy of rheumatic diseases: practice and prospects. Immunol Today 12:177–179, 1991
42. Trentham DE: Immunotherapy and other novel therapies. Curr Opin Rheumatol 3:369–372, 1991

The EMBO Journal vol.12 no.2 pp.725–734, 1993

C10

Human anti-self antibodies with high specificity from phage display libraries

Andrew D.Griffiths[1], Magnus Malmqvist[1], James D.Marks[1], Jacqueline M.Bye[1,2], M.J.Embleton[1], John McCafferty[3], Michael Baier[1], K.Philipp Holliger[1], Barbara D.Gorick[2], Nevin C.Hughes-Jones[2], Hennie R.Hoogenboom[1] and Greg Winter[1,4,5]

[1]MRC Centre for Protein Engineering, [2]MRC Molecular Immunology Unit, MRC Centre, Hills Road, Cambridge CB2 2QH, UK and Division of Transfusion Medicine, University of Cambridge, Cambridge CB2 2PT, [3]Cambridge Antibody Technology Ltd, The Science Park, Melbourn, Cambridgeshire SG8 6EJ and [4]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK
[5]Corresponding author

Communicated by G.Winter

Recently we demonstrated that human antibody fragments with binding activities against foreign antigens can be isolated from repertoires of rearranged V-genes derived from the mRNA of peripheral blood lymphocytes (PBLs) from unimmunized humans. The heavy and light chain V-genes were shuffled at random and cloned for display as single-chain Fv (scFv) fragments on the surface of filamentous phage, and the fragments selected by binding of the phage to antigen. Here we show that from the same phage library we can make scFv fragments encoded by both unmutated and mutated V-genes, with high specificities of binding to human self-antigens. Several of the affinity purified scFv fragments were shown to be a mixture of monomers and dimers in solution by FPLC gel filtration and the binding kinetics of the dimers were determined using surface plasmon resonance ($k_{on} = 10^5-10^6\ M^{-1}s^{-1}$, $k_{off} = 10^{-2}\ s^{-1}$ and $K_a = 10^7\ M^{-1}$). The kinetics of association are typical of known Ab–protein interactions, but the kinetics of dissociation are relatively fast. For therapeutic application, the binding affinities of such antibodies could be improved *in vitro* by mutation and selection for slower dissociation kinetics.

Key words: human antibodies/phage display/self

Introduction

Human monoclonal antibodies (mAbs) have huge potential for therapy, but are difficult to make by immortalizing B-lymphocytes (for reviews see James and Bell, 1987; Winter and Milstein, 1991). Furthermore, it is especially difficult to generate human mAbs directed against human antigens (anti-self antibodies), for example antibodies against soluble TNF to block septic shock (Spooner *et al.*, 1992), against membrane bound carcinoembryonic antigen to image colorectal carcinoma (Mach *et al.*, 1980) or against lymphocyte antigens to destroy tumour in lymphoma (Hale *et al.*, 1988). This difficulty results from immunological tolerance mechanisms that prevent the antigen-driven expansion of B-cell clones with self specificities (Burnet, 1959; Nossal, 1989). After antibody gene rearrangement, virgin B-cells may display antibodies with self-reactivity, but tolerance mechanisms can lead to their deletion (Nossal, 1989; Nemazee *et al.*, 1991; Russell *et al.*, 1991) or to their anergy (Nossal, 1989; Basten *et al.*, 1991; Erikson *et al.*, 1991). It has been suggested that cells may be anergized if the antigen is soluble, but deleted if the antigen is membrane bound (Hartley *et al.*, 1991). B-cell tolerance does not seem to occur when concentrations of soluble antigen are low (in contrast to T-cell tolerance) and B-cells with poor affinities for antigen are not tolerized, even at higher antigen concentrations (Adelstein *et al.*, 1991). Such non-tolerized B-cells are not usually expanded because they lack T-cell help (Bretscher and Cohn, 1970; Adelstein *et al.*, 1991), although proliferation can be induced artificially by using polyclonal B-cell activators (reviewed in Nossal, 1987).

It is estimated that 10–30% of B-lymphocytes in normal, healthy individuals are engaged in making autoantibodies (Cohen and Cooke, 1986). However, the 'natural autoantibodies' produced do not lend themselves to therapeutic use as they are often IgM, low affinity and polyreactive (see Nakamura *et al.*, 1988; Tomer and Schoenfeld, 1988; Casali and Notkins, 1989; Rossi *et al.*, 1990; Avrameas, 1991). An immune response against self can arise in autoimmune disease (see Smith and Steinberg, 1983) or after infections (see Bona, 1988) and a few human mAbs directed against self-antigens have been isolated from patients with active autoimmune disease (see James and Bell, 1987). These autoantibodies are frequently specific, but may bind to only a restricted range of epitopes on the antigen (see Bouanani *et al.*, 1991).

Recently monoclonal antibody fragments have been generated and expressed in bacteria using phage antibody technology (McCafferty *et al.*, 1990) by cloning repertoires of V-genes into filamentous bacteriophage and selecting the recombinant phage with antigen (for review, see Hoogenboom *et al.*, 1992). The repertoires comprised random combinatorial libraries (Huse *et al.*, 1989) of the rearranged heavy and light chain V-genes of immunized animals or human donors. Immunization leads to clonal expansion and production of mRNA by plasma cells: as a result, derived V-gene repertoires are enriched for sequences of heavy and light chains encoding part of an antigen binding site (Hawkins and Winter, 1992). The selected antibody fragments can have good affinities for antigen, for example at least $10^8\ M^{-1}$ for the hapten phOx (Clackson *et al.*, 1991). However, because it is difficult to raise an immune response to self-antigens, we have sought to extend the technology to the generation of human antibodies without the use of immunization.

In principle, a range of binding specificities could be isolated from a single huge and diverse phage library by selection with either self or foreign antigens (for review, see

Table I. Frequency of binding clones isolated from the unimmunized scFv library after selection

Antigen	Rounds of selection					Number of unique clones
	1	2	3	4	5	
Thyroglobulin (bovine)	–	–	18/40	–	–	12
Thyroglobulin (human): selected on bovine	–	–	10/40	–	–	4
Fog1 (human IgGlκ antibody)	–	–	–	94/96	–	4
TNFα (human)	–	122/1920	83/192	92/96	–	7
CEA (human)		–	0/96	1/96	2/96	1
MUC1 (human): selected with peptide	–	–	–	0/96	2/96	1
rsCD4 (human)	–	–	–	–	8/96	1

The ratios indicate the frequency of binding clones after each round of selection. Phagemids were rescued with M13ΔgIII helper phage, except for the CEA, MUC1 and rsCD4 selections, where VCS-M13 helper phage was used.

Marks *et al.*, 1992a). To this end, from a large phage antibody library, we first isolated antibody fragments with a high specificity of binding to foreign antigens (turkey lysozyme, bovine serum albumin and the hapten phenyloxazolone) (Marks *et al.*, 1991) and with affinities (K_a) in the range 10^6-10^7 M^{-1}. Repertoires of rearranged heavy and light chain V-genes were provided by PCR amplification from the μ, κ and λ mRNA of peripheral blood lymphocytes from unimmunized, healthy human donors. The V-genes were assembled (Clackson *et al.*, 1991) at random (Huse *et al.*, 1989) to encode repertoires of single-chain Fv (scFv) fragments (Bird *et al.*, 1988; Huston *et al.*, 1988). The fragments were displayed on the surface of the filamentous bacteriophage (McCafferty *et al.*, 1990) by fusion to the minor coat protein pIII (Smith, 1985), and phage encoding scFv fragments with binding activities were selected by binding of the phage to antigen. On infection of bacteria with the selected phage (Hoogenboom *et al.*, 1991), soluble scFv fragments produced from individual clones by secretion into the bacterial periplasm (Glockshuber *et al.*, 1990) were screened for binding activity. We now demonstrate the use of the same phage library to isolate antibody fragments with high specificity against self-antigens.

Results

The selected human antibody fragments show high specificity against self-antigens

The unimmunized library was subjected to affinity enrichment on a range of antigens (see Materials and methods and Table I). After 2–5 rounds of selection, *Escherichia coli* cells were infected with eluted phage and antibody fragments produced by individual clones were screened for binding by ELISA. Phage selected with the 20 amino acid MUC1 peptide (Price *et al.*, 1990), which corresponds to a repeated motif in human MUC1 mucin (tumour-associated polymorphic epithelial mucin or PEM) (Gendler *et al.*, 1988; Gum *et al.*, 1990), were screened for binding to human PEM and hence bind to both peptide and the protein. The V-genes of clones with binding activities were sequenced and between 1–12 different clones identified for each antigen (Table I). The appearance of only low numbers of clones binding to CEA, PEM and human recombinant soluble CD4 (rsCD4), even after several rounds of selection, may reflect the use of VCS-M13 as helper phage (instead of M13ΔgIII helper used for the other antigens). Populations of phage(mid) particles produced by rescue with M13ΔgIII (which cannot produce pIII) have higher average avidities than those produced by rescue with VCS-M13 (where the wild-type pIII encoded by the helper phage can compete with scFv–pIII fusions).

The scFv fragments were then screened for binding to a panel of other protein antigens and were found to be highly specific. This is illustrated in Figure 1 with the three clones with strongest ELISA signals for bovine thyroglobulin, human TNFα and the human mAb Fog-1, and in Figure 2 with the single clones with binding activity to human CEA, MUC1 and human rsCD4. However for a few clones with poor ELISA signals on the target antigen, we found signals with some of the other proteins of the panel (not shown).

The antibody fragments are derived from a range of unmutated and somatically mutated V-genes

The sequences of several clones with self-specificity are given in Table II and contain both kappa and lambda light chains (Table III). Comparison with the sequences of the nearest germline V-gene segments indicates that several different families are used (VH1, 3, 4 and 5; Vκ1 and 4, Vλ1, 2 and 3). In a few cases the V-genes are completely germline, for example both the VH and Vλ genes of αThy-29. However, most of the V-genes have several differences from the nearest germline V-gene segments, both at the nucleotide and amino acid level (Table III), suggesting that they are derived from somatically mutated B-cells (Berek and Milstein, 1987). Some mutations may have arisen during the PCR amplification and assembly process, for example the VH-genes of αFOG1-G8 and αMUC1-1, and the Vκ-gene of αThy-33 probably arose from cross-overs between two V-genes during PCR amplification (Table III). Furthermore, large differences (for example the Vκ of αFOG1-H6, which differs by 36 nucleotides) may be due to the use of unknown V-gene segments. There is a striking homology in the CDR3 of the heavy chain between αTNF-A1 and αTNF-E1: the germline V-genes are different but the same JH segments are used and 11 out of 16 residues of CDR3 are identical. This suggests that both scFv fragments may bind to the same epitope of TNF.

The antibody fragments are directed to different epitopes on the same protein

The scFv fragments directed against bovine thyroglobulin were screened for binding to human thyroglobulin, which differs by only six single amino acid residues in the protomer (Malthiéry and Lissitzky, 1987). Four of the twelve clones (including αThy-29) bound to human thyroglobulin, whereas

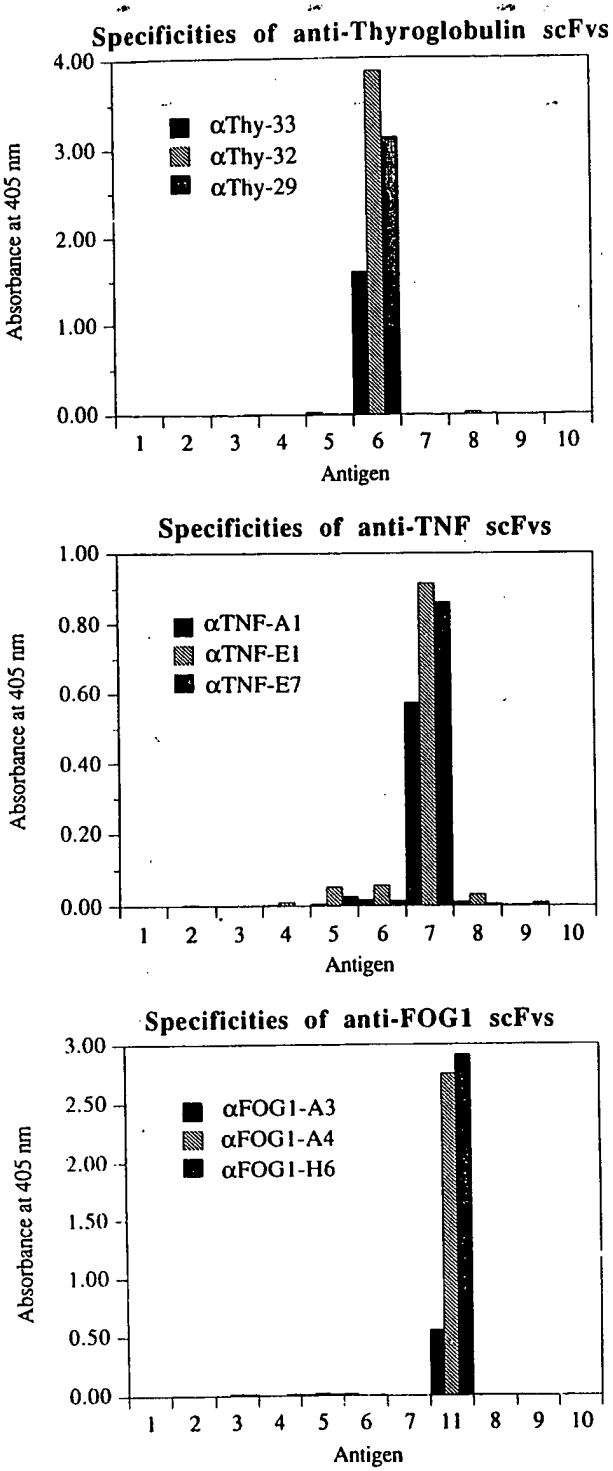


Fig. 1. Specificities of soluble single-chain Fvs (scFvs) isolated from the unimmunized library by selection on bovine thyroglobulin **(upper panel)**, human TNFα **(centre panel)** or the human mAb Fog-1 ($\gamma 1, \kappa$) (Melamed *et al.*, 1987) **(lower panel)**. Binding was determined by ELISA to a panel of proteins: 1, plastic; 2, hen egg trypsin inhibitor; 3, chymotrypsinogen A; 4, hen egg ovalbumin; 5, keyhole limpet haemocyanin; 6, bovine thyroglobulin; 7, human TNFα; 8, turkey egg-white lysozyme; 9, horse heart cytochrome c; 10, bovine serum albumin; 11, mAb Fog-1.

the rest (including αThy-32 and αThy-33) did not (data not shown). Likewise the fragments binding to the human antibody Fog-1 were screened for binding to a range of other antibodies differing in heavy and light chain isotype (Figure 3). The fragment αFOG1-A4 bound to all heavy

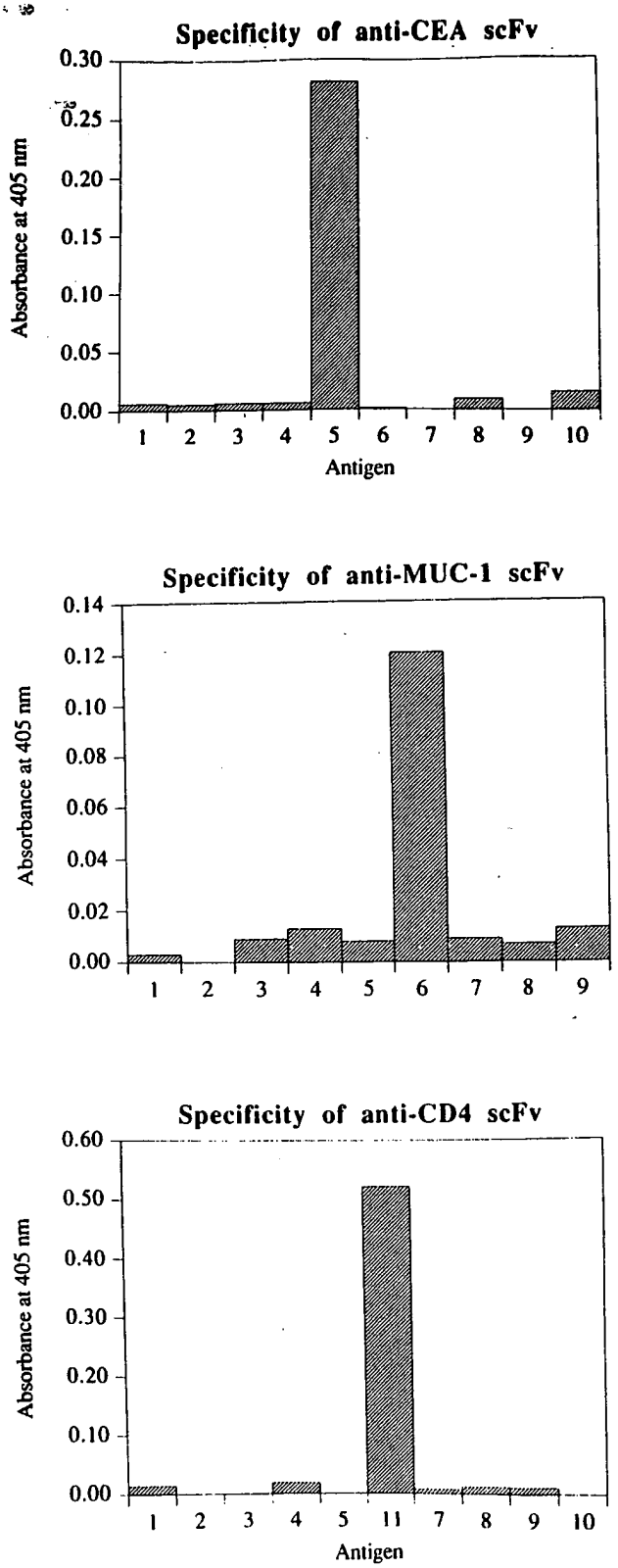


Fig. 2. Specificities of soluble single-chain Fvs (scFvs) isolated from the unimmunized library by selection on human CEA **(upper panel)**, the MUC1 peptide (Price *et al.*, 1990) **(central panel)** or human CD4 **(lower panel)**. Binding was determined by ELISA to a panel of proteins: 1, hen egg trypsin inhibitor; 2, chymotrypsinogen A; 3, hen egg ovalbumin; 4, keyhole limpet haemocyanin; 5, CEA; 6, urine extract containing human polymorphic epithelial mucin (PEM); 7, bovine thyroglobulin; 8, hen egg-white lysozyme; 9, bovine serum albumin; 10, chicken γ globulin coupled to 4-hydroxy-3-nitrophenyl acetic acid; 11, human recombinant soluble CD4.

Table II. Deduced protein sequences of several antigen-specific scFv fragments isolated from the unimmunized library

A. Heavy chains

scFv	FR 1	CDR 1	FR 2	CDR 2	FR 3	CDR 3	FR 4
αThy-23	QVQLQQSGGGVVQPGRSMRLSCAASGFNFR	SYGMH	WVRQAPGKGLEWVS	GISGSGGSTYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK	GSMIVVARYFDY	WGQGTLVTVSS
αThy-29	QVQLVQSGAEVKKPGASVKVSCKASGYTFT	SYGIS	WVRQAPGQGLEWMG	WISAYNGNTNYAQKLQG	RVTMTTDTSTSTAYMELRSLRSDDTAVYYCAA	DTGRIDDFWSGYNFDY	WGQGTLVT
αThy-32	QVQLVQSGGGLVQPGRSLRLSCAASGFTFD	DYAMH	WVRQAPGKGLEWVS	GISSNSGSIGYADSVKG	RFTISRDNAKNSLYLQMNSLRAEDTALYYCAR	GIAVAGAYYFDY	WGQGTLVTVSS
αThy-33	QVQLVQSGGGVVQPGGSLRLSCAASGLSIR	TNGMH	WVRQAPGKGLEWVA	AISYDGRSVYYADSVQG	RVTISRDNSKNTVHLQITSLKSEDTAVYYCAK	DSSSSWFLDS	WGQGTLV
αFOG1-A3	EVQLVESGGGLVQPGGSLRLSCAASGFTFS	SYWMS	WVRQAPGKGLEWVA	NIKQDGSEKYYVDSVKG	RFTISRDNAKNTLYLQMNSLRDEDTAVYYCAK	NPRDSGSYYYFDY	WGQ
αFOG1-A4	QVQLQESGGSVVQPGRSLRLSCAASGFTFS	NYAIH	WVRQAPGKGLEWVA	VISYDGGYEYYADSVKG	RSTISRDNSKNTLYLQMNSLRAEDTAVYYCAR	DASVHTAPYYYMDV	WGK
αFOG1-H6	QVQLQQSGGSLVQPGGSLRLSCAASGFTFS	NSGMN	WVRQAPGKGLEWVS	YISSSSSTIYYADSVKG	RFTISRDNAKNSLYLQMNSLRDEDTAVYYCAR	EEGGLMDV	WGKG
αFOG1-G8	QVQLQESGAGLLKPSETLSLTCAVYGGSFS	GYYWIG	WVRQMPGKGLEWMG	IINPGDSDTRYSPSFQG	QVTISVDKSVSTAYLQWSSLKPSDSAVYFCAR	HDVGYCSSPNCAKRPEYFQH	WGQ
αTNF-A1	EVQLVESGGGLVQPGGSLRLSCAASGFTFS	SYGMH	WVLQAPGKGLEWVA	FIRYDGSNKYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCVR	EDHVITTGRYHYYMDV	WGK
αTNF-E1	QVQLQESGGGLVQPGGSLRLSCAASGLTFS	SYAMH	WVRQAPGKGLEWVA	VISYDGSNKYYADSVKG	RFTISRDNSKNTLYLQMDNLRAEDTAVYYCVR	EDYVITSGFYYYYMDV	WGK
αTNF-E7	QVQLQESGAEVKKPGSSVKVSCKASGGTFS	SYAIS	WVRQAPGQGLEWMG	GIIPIFGTANYAQKFQG	RVTITADESTSTAYMELSSLRSEDTAVYYCAR	GPLRGYDYYYYYMDV	WGK
αTNF-H9G1	QVQLVQSGAEVKKTGSSVKVSCKTSGYTFT	YRYLH	WVRQAPGQALEWMG	WITPFNGNTNYAQKFQD	RVTITRDRSMSTAYMELSSLRSEDTAMYYCAR	SGDLYSGYED	WGQ
αCEA4-8A	QVQLQQSGAEVKKPGASVKVSCKASGYTFT	SYGIS	WVRQAPGQGLEWMG	WISAYNGNTNYAQKLQG	RVTMTTDTSTSTAYMELRSLRSDDTAVYYCAR	DSFGYCSSTSCPYYYYYMDV	WGKGTTVTVSS
αMUC1-1	QVQLVQSGAEVKKPGASVKVSCKASGYTFT	GYYMH	WVRQAPGQGLEWMG	WINPNSGGTNYAQKFQG	RVTITRDTSASTAYMELSSLRSEDTAVYYCAR	DFLSGYLDY	WGQGTLVTVSS
αCD4-74	QVQLQQSGAEMKKPGESLKISCKGFGYDFS	TYWIG	WVRQMPGKGLEYMG	LIYPGDSDTKYSPSFQG	QVTISADKSISTAYLQWSSLKASDTAMYYCAR	VSGYCSSTSSYDYYYYYMDV	WG

B. Light chains

scFv	FR 1	CDR 1	FR 2	CDR 2	FR 3	CDR 3	FR 4
αThy-23	DIQMTQSPSSLSASVGDSVTITC	QASQGIRNDLA	WYQQKPGKAPKLLIY	AASTLQS	GVPSRFSGSGSGTEFTLTINGLQPEDFATYYC	QQLGAYPLT	FGGGTKLEIKR
αThy-29	SSELTQDPAVSVALGQTVRITC	QGDSLRSYYAS	WYQQKPGQAPVLVIY	GKNNRPS	GIPDRFSGSSSGNTASLTITGAQAEDEADYYC	NSRDSSGNLYVV	FGGGTKLTVLG
αThy-32	QSVLTQPPSVSGAPGQRVTISC	TGSSNIGAGYDVH	WYQQLPGAAPKLLIY	GNSNRPS	GVPDRFSGSKSGTSASLAITGLQAEDEADYYC	QSYDSSLSGWV	FGGGTKVTVLG
αThy-33	DVVMTQSPSTVSASVGDRVTLTC	RASQSISRWLA	WYQQKPGQPPKLLIY	WASTRKS	GVPDRFSGSGSGTDFTLTISSLQADDFATYYC	QHYDSFSPS	FGQGTKVEIK
αFOG1-A3	QSALTQPASVSGSPGQSITISC	TGTSSDVGGYNYVS	WYQQYPGKAPKLLIY	EVSKRPS	GVSNRFSGSKSGNTASLTISGLQTEDEADYFC	SAYAPTGIMM	FGGGTKLTVLG
αFOG1-A4	DIQMTQSPSSLSASVGDRVTITC	RASQGISNYLA	WYQQKPGKVPKLLIY	AASTLQS	GVPSRFSGSGSGTDFTLTISSLQPEDVAVYYC	QQYYSTPT	FGQGTKVEIKR
αFOG1-H6	DIQMTQSPSTLSASIGDRVTITC	RASQNIDSWLA	WFQHKPGKAPKPLIY	GAFTLQN	GVPSRFSGSGSGTEFSLTISSLQPDDFATYFC	QQAHSFPPT	FGGGTRLEIKR
αFOG1-G8	DIQMTQSPSTLSASIGDRVTITC	RASQSIGLWLA	WYQQKPGKAPKLLIY	AASTLQT	GVPSTFSGSGSGTEFTLTISSLRPEDFATYYC	QQLISYPLT	FGGGTKVEIKR
αTNF-A1	DIQMTQSPSSLSASVGDRVTITC	RESQGIRNDLG	WYQQKPGKAPKLLIY	GTSSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQTTSFPLT	FGGGTKLEIK
αTNF-E1	EIVLTQSPSSVSASVGDRVTITC	RASQGIRSWLA	WYQQKPGKAPNLLIY	AASTLET	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQANSFPLT	FGGGTKLE
αTNF-E7	DIVMTQSPSSLSASIGDRVTITC	RESQGIRNDLG	WYQQKPGKAPKLLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC	QQANSLPPT	FGGGTKVEIKR
αTNF-H9G1	QSVLTQPPSVSGPPGQRVTISC	TGTSSNIGAGYDVY	WYQQLPDTSPRVLIY	RNSNRPS	GVPDRFSGSKSGTSASLAITGLQAEDEADYYC	QSYDSRLIRV	FGGGTKLTVLG
αCEA4-8A	EIVLTQSPSSLSASVGDRVTITC	RASQSISSYLN	WYQQKPGKAPKLLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQYSNYPLT	FGGGTKVDIK
αMUC1-1	QSVLTQPASVSGSPDQSITISC	TGTSSDVGGYKYVS	WYQQHPDKAPKVMIY	DVTNRPS	GGSNRFSGSKSGNTASLTISGLQAEDEADYYC	SSYAGAQSLV	FGGGTKLTVLG
αCD4-74	HVILTQPPSASVAPGQRVTISC	SGSRSNIGSNPVS	WYRQFPGAAPKLLIY	IDDRRSS	GVPDRVSGSRSGTSASLAISGLQSEDEADYYC	VAWDDSLKGWV	FGEGTKLTVL

FR, framework region; CDR, complementarity-determining region. Bovine thyroglobulin binders; αThy-23, αThy-29, αThy-32 and αThy-33. Human thyroglobulin binders; αThy-23 and αThy-29. Fog-1 (a human γ1/κ mAb) binders; αFOG1-A3, αFOG1-A4, αFOG1-H6 and αFOG1-G8. Human TNFα binders; αTNF-A1, αTNF-E1, αTNF-E7 and αTNF-H9G1. Human CEA binder; αCEA4-8A. Human MUC1-1 binder; αMUC1-1. Human rsCD4 binder; αCD4-74.

chain γ1, 2 and 3 isotypes, but not to γ4 or μ. By contrast, the fragments αFOG1-H6 and αFOG1-A3 did not bind to any of the other antibodies, including those of the same isotype as Fog-1, suggesting that they are directed to the variable domain of Fog-1.

Two of the antibody fragments are directed against idiotopes of human mAb Fòg-1

The binding of ^{125}I-Fog-1 antibody to human red blood cells bearing the Rh D antigen could be inhibited by both αFOG1-H6 and αFOG1-A3 scFv fragments. Hence, both αFOG1-H6 and αFOG1-A3 are site-associated anti-idiotype antibodies, complexing with the antigen-binding site of Fog-1. The extent of inhibition of ^{125}I-Fog-1 binding to the Rh D antigen (on human R_1R_2 red blood cells) was determined by titration with affinity purified αFOG1-H6 and αFOG1-A3 scFv fragments. [As control, no inhibition of ^{125}I-Fog-1 binding was observed using a scFv fragment (αTEL9) (Marks *et al.*, 1991) directed against turkey egg-white lysozyme.] With the maximum of 16 μg scFv (1000-fold molar excess to ^{125}I-Fog-1), the binding was inhibited by 14.2% (αFOG1-H6) and 20.9% (αFOG1-A3), suggesting that the affinities of these fragments for Fog-1 are much lower than the affinity of Fog-1 for the Rh D antigen ($K_a = 2.2\times10^9$ M^{-1}) which binds monovalently (Gorick *et al.*, 1988). If 100% of the fragments are active, the affinities of the two fragments for binding to Fog-1 could be estimated as $K_a = 3\times10^5$ M^{-1} for αFOG1-H6 and 6×10^5 M^{-1} for αFOG1-A3 and this is consistent with other kinetic measurements (see below and Table IV).

The scFv fragments can form both monomers and dimers in solution

Soluble antibody fragments were purified from bacterial supernatants by affinity chromatography, by binding of the C-terminal peptide tag to the mAb 9E10 (Munro and Pelham, 1986; Clackson *et al.*, 1991; Marks *et al.*, 1991). After ultrafiltration, the fragments were further purified by FPLC gel filtration (Pharmacia) on Superdex 75 (Pharmacia), and detected on-line both by UV absorption (280 nm) and by binding to antigen immobilized on a sensor chip in BIAcore (Pharmacia Biosensor AB) (Jönsson *et al.*, 1991; Jönsson and Malmqvist, 1992). This showed that the scFv fragments emerged in two peaks, corresponding in size to monomers and dimers (Figure 4). The dimers bind more strongly to the immobilized antigen than monomers due to their greater avidity of binding. The scFv dimers run as monomers on non-reducing SDS gels (Laemmli, 1970) (not shown) and are therefore not linked by disulphide bonds. As two peaks are seen in gel filtration, it appears that in this case the monomers and dimers do not interconvert rapidly (for discussion and references for gel filtration equilibria see Jones *et al.*, 1985). Presumably the dimers are scFv fragments interlocked through the flexible linker joining the heavy and light chains, or with the heavy chain of one scFv molecule associated with the light chain of the other. We have preliminary evidence that antibody Fab fragments made in bacteria can also multimerize (unpublished data).

The scFv fragments have micromolar affinities

The presence of both scFv monomers and dimers could lead to an overestimate of affinity of binding using solid phase

Table III. V-gene family, germline derivation and extent of somatic hypermutation of several antigen-specific scFv fragments isolated from the unimmunized library

scFv	Family	Germline genes of closest nucleotide sequence	Differences from germline	
			Nucleotide	Aminoacid
Heavy chains				
αThy-23	VH3	DP-47	13	8
αThy-29	VH1	DP-14	0	0
αThy-32	VH3	DP-31	5	2
αThy-33	VH3	DP-49	32	19
αFOG1-A3	VH3	DP-54	7	3
αFOG1-A4	VH3	DP-46	7	7
αFOG1-H6	VH3	DP-51	10	4
αFOG1-G8[a]	VH4	DP-63 (FR1)	2	0
	VH5	DP-73 (CDR1 to FR3)	15	7
αTNF-A1	VH3	DP-50	9	6
αTNF-E1	VH3	DP-46	14	6
αTNF-E7	VH1	DP-10	0	0
αTNF-H9G1	VH1	DP-4	1	1
αCEA4-8A	VH1	DP14	1	0
αMUC1-1[a]	VH1	VI-2 (FR1 to CDR2)	2	0
	VH1	DP-25 (FR3)	0	0
αCD4-74	VH5	DP-73	13	8
Light chains				
αThy-23	Vκ1	L8	20	9
αThy-29	Vλ3	IGLV3S1	0	0
αThy-32	Vλ1	IGLV1S2	1	1
αThy-33[a]	Vκ1	L12 (FR1 and CDR1)	6	3
	Vκ4	B3 (FR2 to FR3)	5	5
αFOG1-A3	Vλ2	VL2.1	16	9
αFOG1-A4	Vκ1	O4	25	12
αFOG1-H6	Vκ1	L5	36	17
αFOG1-G8	Vκ1	L8	25	10
αTNF-A1	Vκ1	L11	12	8
αTNF-E1	Vκ1	L5	5	5
αTNF-E7	Vκ1	L11	17	8
αTNF-H9G1	Vλ1	IGLV1S2	18	9
αCEA4-8A	Vκ1	O2	4	0
αMUC1-1	Vλ2	VL2.1	18	12
αCD4-74	Vλ1	Humlv1L1	23	17

References for all the heavy chain germline genes can be found in Tomlinson *et al.* (1992). The references for the light chains are VL2.1 (Brockly *et al.*, 1989); IGLV1S2 (Bernard *et al.*, 1990); IGLV3S1 (Frippiat *et al.*, 1990); L8(Vd) and L5(Vb)(Pech *et al.*, 1984); L12(HK102) (Bentley and Rabbits, 1980); B3(VKIV) (Klobeck *et al.*, 1985); 02 and 04 (Pargent *et al.*, 1991); L11 (Scott *et al.*, 1991); Humlv1L1 (Daley *et al.*, 1992). Alternative names are given in parenthesis. a) These genes appear to have been created by cross-overs between two V-genes during PCR amplification and therefore matches have been determined using the two putative germline segments: FR, framework; CDR, complementarity-determining region.

methods. To determine the affinity and kinetics of binding of scFv fragments to the antigen-coated chip using surface plasmon resonance, we therefore purified the fragments by gel filtration (Table IV). For the dimers, the off-rate constants were determined as $\sim 10^{-2}$ s^{-1} and the on-rate constants for the scFv dimers as $\sim 10^5 - 10^6$ M^{-1} s^{-1} (assuming the sample is completely active). In the case of αFOG1-H6, the antigen (the mAb Fog-1) was immobilized on the sensor chip in two ways, either directly (Figure 5) or via a rabbit anti-mouse IgG1 antibody. The results were almost identical by either method (see Table IV). However the active fraction of scFv fragments varies considerably and could lead to an underestimate of the on-rate (and affinity of binding); for example using fluorescence quench titration with several scFv fragments directed against phenyloxazolone we detected only 0.06−0.38 functional binding sites per scFv molecule (unpublished data). Indeed the on-rate constants calculated for the association of the αFOG1-H6 fragment and Fog-1 antibody depend on whether the antibody (k_{on} 2.2×10^5 M^{-1} s^{-1}) or scFv fragment (k_{on} 1.0×10^6 M^{-1} s^{-1}) is immobilized on the sensor chip (Table IV), indicating that the αFOG1-H6 fragment is less active than the Fog-1 antibody. For the scFv monomers, the binding signals were low and it was difficult to follow the kinetics of binding to the surface, except for the dissociation of the αThy-29 monomer ($k_{off} = 2\times10^{-2}$ s^{-1}). However, the 4-fold stabilization of the αThy-29 fragment dimer (see below), suggests that the off-rate constants of the other monomers are $> 10^{-2}$ s^{-1}, perhaps 10^{-1} s^{-1}.

The greater stability of the scFv dimers on the sensor chip

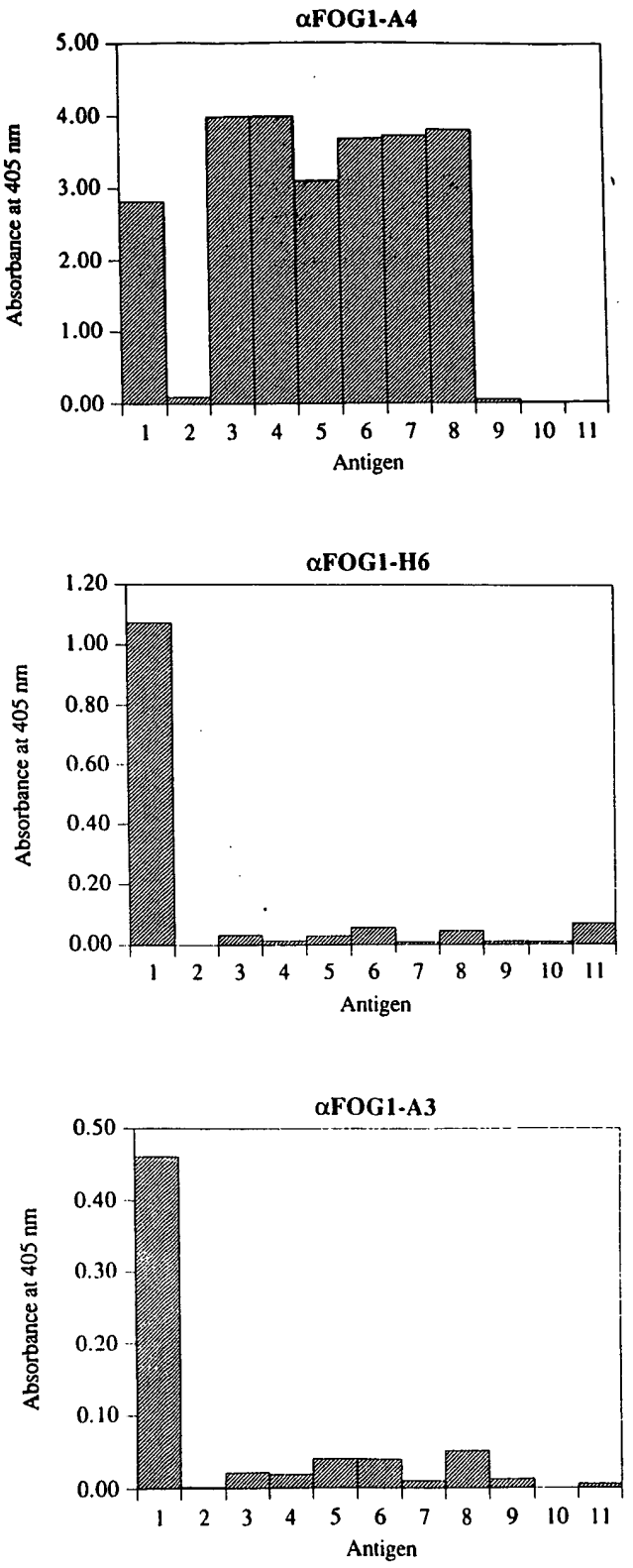


Fig. 3. ELISA to assay the binding of three scFvs, isolated by selection on a human monoclonal antibody Fog-1 (IgG1, κ) (Melamed *et al.*, 1987), to a panel of human antibodies of varying isotype: 1, Fog-1; 2, the Fv fragment of Hulys11 (Foote and Winter, 1992); 3, Hulys11 antibody (IgG1, κ); 4, RegA (IgG1, κ) (Melamed *et al.*, 1987); 5, FogC (IgG3, κ) (N.C.Hughes-Jones, unpublished); 6, Pag1 (IgG1, λ) (Thompson *et al.*, 1986); 7, IgG2, λ antibody purified from myeloma plasma (Sigma); 8, Oak3, (IgG3, λ) (Bye *et al.*, 1992); 9, IgG4, λ purified from myeloma plasma (Sigma); 10, Foml (IgM, λ) (Melamed *et al.*, 1987); 11, FomA (IgM, λ) (Melamed *et al.*, 1987).

compared with monomers indicates that the dimers are bivalent. The scFv dimers are therefore analogous to the two heads of the antibody IgG, but with different spacing between the heads, and their binding avidities were estimated as $\sim 10^7$ M^{-1} from k_{on}/k_{off} (Table IV). The affinities of the monomers must be lower by virtue of their faster dissociation from the surface. For the αThy-29 monomer, and assuming that the on-rate constant is the same as for the dimer (Mason and Williams, 1986), we can estimate an affinity of $\sim 3\times10^6$ M^{-1}. These affinities, calculated from the rate constants measured by surface plasmon resonance, appear to be similar to those measured in solution by fluorescence quench techniques. For example the affinity of binding of the monomer scFv fragment αTEL9 (Marks *et al.*, 1991), which binds to turkey lysozyme (and was derived from the same library), was estimated as 3.9×10^7 M^{-1} using surface plasmon resonance (Table IV), and as 1.2×10^7 M^{-1} by fluorescence quench (Marks *et al.*, 1991).

Discussion

We had demonstrated previously that highly specific human antibody fragments (scFv), directed against 'foreign' antigens, both protein and hapten, could be isolated from a large phage display library composed of the rearranged V-genes of peripheral blood lymphocytes from unimmunized donors (Marks *et al.*, 1991). Here we have shown that antibody fragments directed against human antigens (self-antigens), including idiotopes of a human antibody, a cytokine (TNFα), two tumour markers (CEA and MUC1) and the T-lymphocyte marker CD4 can be derived from the same library. Immunological tolerance would make immunization of humans with these antigens difficult; TNFα is also extremely toxic and raising an immune response against CD4 would be an act of suicide by the immune system. The antibody fragments we have isolated show a high specificity of binding to antigen. This contrasts with the poor specificity of binding of fragments isolated from a phage display library in which mouse Fab fragments were fused to the major coat protein (pVIII) of filamentous phage (Gram *et al.*, 1992). The use of different V-genes and heavy and light chain combinations for each antigen suggested that each fragment was likely to bind to different epitopes and this was shown directly for the fragments against thyroglobulin and the human mAb Fog-1.

The affinity of antibodies isolated from a library is thought to be proportional to the library size (Perelson and Oster, 1979) and in this case the size of the phage library is comparable to the number of B-cells in a mouse, and the affinities of antibodies isolated are typical of antibodies from the mouse primary immune response (Foote and Milstein, 1991). The kinetics of association of the antibody fragments to the protein self-antigens (10^5-10^6 M^{-1} s^{-1}) are also typical of previously characterized Ab–protein interactions. However the kinetics of dissociation (10^{-2} s^{-1}) are relatively fast for Ab–protein interactions (but both rates are slow compared with many Ab–hapten interactions) (Smith and Skubitz, 1975; Pecht, 1982; Mason and Williams, 1986; Foote and Milstein, 1991; Foote and Winter, 1992). At first sight, it is surprising that we can isolate scFv fragments with such fast off-rates, as a 'monomeric' phage should not be retained on the solid support during washing. However, scFv fragments are

Table IV. Affinities and kinetics of antigen binding by monomeric and dimeric scFv fragments

scFv	(M/D)[a]	Immobilized species	k_{on}[b] (BIAcore) $M^{-1}s^{-1}/10^4$	k_{off}[b](BIAcore) $s^{-1}/10^{-2}$	$K_a=k_{on}/k_{off}$ (BIAcore) $M^{-1}/10^6$	K_a by FQ[c] or inhibition[d] $M^{-1}/10^6$
αTNF-E7	D	HumanTNFα	9.0 (±1.2)	1.4 (±0.054)	6.4	ND
αFOG1-H6	D	Fog-1 (direct)	22.2 (±0.4)	1.8 (±0.23)	12.3	ND
αFOG1-H6	D	Fog-1 (via RAMIgG1)	22.1 (±1.9)	2.4 (±0.045)	9.3	ND
αFOG1-H6	D	αFOG1-H6 scFv	104 (±2.4)	ND[e]	ND	ND
αFOG1-H6	M + D	(Measured by inhibition)	ND	ND	ND	0.3[d]
αFOG1-A3	M + D	(Measured by inhibition)	ND	ND	ND	0.6[d]
αThy-29	D	Human thyroglobulin	6.6 (±1.2)	0.46 (±0.063)	14.3	ND
αThy-29	M	Human thyroglobulin	ND	2.0 (±0.37)	ND	ND
αTEL9	M	Turkey egg lysozyme	39.2 (±2.6)	1.0 (±0.97)	39.2	11.6[c]

[a] M, monomeric fraction; D, dimeric fraction.
[b] Numbers in brackets are standard deviations.
[c] FQ, fluorescence quench titration.
[d] Calculated from the extent of inhibition of ^{125}I-Fog-1 binding to the Rh D antigen.
[e] Not determined because the dissociation curves were very badly bent.

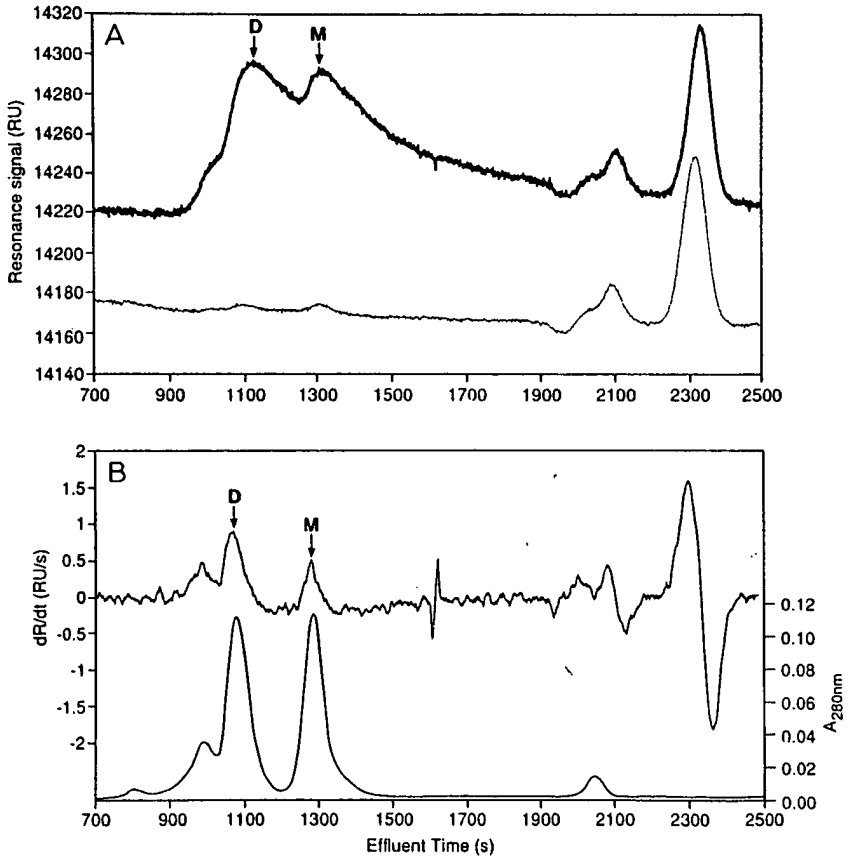


Fig. 4. Gel filtration of affinity purified soluble scFv αThy-29 on Superdex 75 analysed by UV absorption and on-line specific detection of the active component on BIAcore. A. BIAcore sensorgram [resonance signal (RU) as a function of time] showing adsorption of scFv in the column effluent passing over a sensor chip with immobilized human thyroglobulin (thick line) and the same sample run over a bare CM5 sensor chip surface without any antigen (thin line). B. UV profile of the gel filtration (lower line) and the derivatized sensorgram (upper line) which illustrates the rate of change in mass of protein bound to the sensor chip as a function of time. M, scFv monomer; D, scFv dimer.

displayed multivalently on the phage, especially using the M13ΔgIII helper phage, and some of the scFvs that tend to form dimers in solution may also form dimers on phage. The multivalent interactions with antigen help retain the phage, allowing the encoded scFv phage to be isolated.

Random combinatorial V-gene repertoires derived from the mRNA of immunized animals are enriched for heavy or light chain V-genes encoding part of an antigen binding site (Winter and Milstein, 1991; Hawkins and Winter, 1992), and this facilitates the isolation of antigen-binding fragments (Clackson *et al.*, 1991; Persson *et al.*, 1991) although the combinations of V-genes of each B-lymphocyte appear to be largely destroyed (Winter and Milstein, 1991; Gherardi and Milstein, 1992). Antigen binding sites can also be generated *de novo* by the random combination of chains, as illustrated by the isolation of scFv fragments against

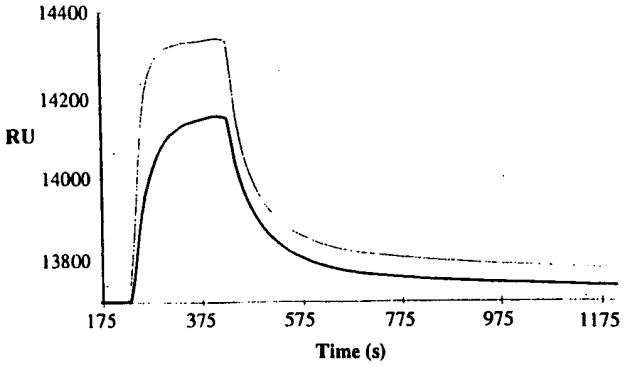


Fig. 5. BIAcore sensorgram [resonance signal (RU) as a function of time] of the interaction of soluble αFOG1-H6 scFv dimer with immobilized mAb Fog-1 (Melamed *et al.*, 1987). A 35 μl pulse of 200 nM (thin line) or 80 nM (thick line) scFv protein was passed, with a flow rate of 10 μl/min over a sensor chip to which mAb Fog-1 was coupled directly.

foreign antigens from unimmunized human donors (Marks *et al.*, 1991). However, the origins of the V-genes of scFv fragments directed against self-antigens are less clear. Self-reactive antibodies, including those with specificities against human thyroglobulin (Ruf *et al.*, 1985), human TNFα (Bendtzen *et al.*, 1990) and human IgG (Welch *et al.*, 1983), are common in healthy individuals and indeed 10–30% of B-lymphocytes appear to be engaged in making autoantibodies (Cohen and Cooke, 1986). Therefore the V-genes could be derived from B-cells that are autoreactive, or those that are not. Since somatic hypermutation of antibody genes is triggered only after antigen-induced B-cell proliferation (Griffiths *et al.*, 1984), the isolation of scFv fragments encoded by somatically mutated V-genes (Table III) indicates that the V-genes have been derived from lymphocytes that have been stimulated by antigen; for example from B-cells with self-specificities that have been stimulated with cross-reactive foreign antigen, or from B-cells encoding antibodies of other (foreign) specificities. Conversely those scFv fragments encoded by V-genes with little or no somatic mutation (see Table III) may well have been derived from virgin B-cells or those involved in early immune responses.

'Natural autoantibodies' (self-reactive antibodies isolated from healthy donors) tend to be of low affinity and polyspecific and may well be produced by a discrete subset of B-cells, the internal activity set (Holmberg and Coutinho, 1985), contributed in part by $CD5^+$ B-cells (Casali and Notkins, 1989). In contrast, these anti-self scFv fragments are highly specific in binding to antigen despite only having micromolar affinities. However, their affinities could presumably be improved *in vitro*, for example, the affinity of an scFv fragment for the hapten phenyloxazolone derived from the phage library (and like the anti-self antibodies described here with a relatively fast off-rate) was improved from $K_a = 3.1\times10^6\ M^{-1}$ to $9.1\times10^8\ M^{-1}$ by chain shuffling (Marks *et al.*, 1992b). This would allow the creation of highly specific, high affinity human antibodies directed against self-antigens for use in human therapy.

Materials and methods

Selection of phage library

The construction and selection of the library of phages displaying scFv fragments has been described previously by Marks *et al.* (1991) and as briefly summarized in the Introduction. The library used in this work contained 2.9×10^7 clones and the VH genes were amplified from cDNA primed with an IgM-specific constant region primer.

To rescue the library, 50 ml of 2×TY broth (Miller, 1972) containing 100 μg ampicillin/ml, 1% glucose (2×TY-AMP-GLU) were inoculated with 10^9 *E.coli* TG1; (Gibson, 1984) of the library stock (~10 μl) and grown, shaking, at 37°C until the culture reached an $OD_{600\ nm}$ of 0.5. 5 ml of this culture (~2.5×10^9 cells) were added to 50 ml 2×TY-AMP-GLU (pre-warmed to 37°C) containing 5×10^{10} p.f.u. VCS-M13 helper phage (Stratagene) or 5×10^8 pfu M13ΔgIII helper phage (unpublished data). M13ΔgIII helper phage does not encode pIII hence the phage(mid) displaying antibody fragments have a greater avidity of binding to antigen. Infectious M13ΔgIII particles were made by growing the helper phage in cells harbouring a pUC19 derivative supplying the wild-type gIII protein during phage morphogenesis. The culture was incubated for 1 h at 37°C without shaking and then a further 1 h at 37°C with shaking. Cells were spun down (IEC-Centra 8, 4000 r.p.m. for 10 min), resuspended in 300 ml 2×TY broth containing 100 μg/ml ampicillin and 25 μg/ml kanamycin (2×TY-AMP-KAN) and grown overnight shaking at 37°C (or 25°C for phage selected with rsCD4). Phage particles were purified and concentrated from the culture medium by two PEG precipitations (Sambrook *et al.*, 1990), resuspended in 2 ml PBS and passed through a 0.45 μM filter (Minisart NML; Sartorius) to give a final concentration of ~10^{13} transducing units/ml (ampicillin-resistant clones).

The phage were panned for binding using immuno tubes (Nunc; Maxisorp) coated with antigen essentially as Marks *et al.* (1991), or were selected on a column of antigen (McCafferty *et al.*, 1990). Six antigens were used: a human mAb Fog-1 (γ1, κ) (Melamed *et al.*, 1987); recombinant human tumour necrosis factor-α (TNFα) (expressed in yeast); bovine thyroglobulin (Sigma); human recombinant soluble CD4 (rsCD4) (expressed in baculovirus by American Biotechnologies Inc. and supplied by the MRC AIDS Reagent Project [ADP608]); human carcinoembryonic antigen (CEA); and a 20 amino acid peptide (Price *et al.*, 1990), which corresponds to a repeated motif in human MUC1 mucin (tumour-associated polymorphic epithelial mucin or PEM) (Gendler *et al.*, 1988; Gum *et al.*, 1990). All antigens, except the MUC1 peptide, were coated on immuno tubes overnight at room temperature. 10 μg/ml TNFα was coated in 50 mM $NaHCO_3$ (pH 9.6), whilst the other antigens were coated in PBS at a concentration of 10 μg/ml Fog1 and rsCD4, 1 mg/ml bovine thyroglobulin or 20 μg/ml CEA. For the first two rounds of selection tubes were washed 10 times with PBS, 0.1% (v/v) Tween 20 and 10 times with PBS. For subsequent rounds of selection tubes were washed 20 times with PBS, 0.1% (v/v) Tween 20 and 20 times with PBS. Phage were eluted with 100 mM triethylamine as Marks *et al.* (1991). Eluted phage (usually 10^6-10^7 transducing units) were used to infect *E.coli* TG1 cells. Approximately 10^9 infected bacteria were used as an inoculum for the next rescue. The library was subjected to 3–5 rounds of rescue and selection for each antigen.

For selection of phage binding to the MUC1 peptide, the peptide was coupled chemically to Sepharose 4B (provided by M.R.Price). A 1 ml column was prepared and phage was selected as described by McCafferty *et al.* (1990). Briefly, the Sepharose–MUC1 column was washed with PBS containing 2% skimmed milk powder (MPBS) and the phage loaded in 1ml of the same buffer. After washing the column successively with 10 ml volumes of MPBS, PBS (pH 7.2), 50 mM Tris–HCl–500 mM NaCl (pH 8.0) and 50 mM Tris–HCl–500 mM NaCl (pH 9.0), phage was eluted with 5 ml 100 mM triethylamine and neutralized with 0.5 M sodium phosphate buffer (pH 6.8). Five rounds of selection were carried out.

Screening and sequencing of clones

Single ampicillin resistant colonies from infection of *E.coli* TG1 with eluted phage, were screened either for binding of phage (Clackson *et al.*, 1991) or soluble scFv fragments (Marks *et al.*, 1991). Since the gene encoding the antibody fragment is linked to that encoding the phage coat protein by an amber codon, soluble fragments can be secreted from a non-suppressor strain of bacteria infected by the phage (Hoogenboom *et al.*, 1991). The binding to antigen of soluble scFvs in bacterial supernatant was detected with the mouse mAb 9E10 (1 μg/ml), which recognizes the C-terminal peptide tag (Munro and Pelham, 1986) and peroxidase-conjugated anti-mouse Fc antibody (Sigma), as described by Ward *et al.* (1989). Plates were coated with the antigens Fog1, TNFα, bovine thyroglobulin and rsCD4 as described for immuno tubes above and with CEA at 5 mg/ml. A urine extract containing human polymorphic epithelial mucin (PEM) was used at a protein concentration of ~10 mg/ml.

The specificity of the isolated clones was checked by ELISA of the soluble scFv fragments using plates coated with various proteins. Plates were coated with the antigens Fog-1, TNFα, bovine thyroglobulin, rsCD4, CEA and PEM as described above. Other proteins were coated overnight at room temperature at a concentration of 1 mg/ml in PBS (cytochrome *c* [Sigma]) or in 50 mM $NaHCO_3$ (pH 9.6) and bovine serum albumin, turkey egg-

white lysozyme, hen egg-white lysozyme, hen ovalbumin, keyhole limpet haemocyanin (CalBiochem), chymotrypsinogen A, chicken egg-white trypsin inhibitor (Sigma) and chicken γ globulin coupled to 4-hydroxy-3-nitrophenyl acetic acid. The Fog-1 specific clones were screened by binding to a panel of different human antibodies (see legend to Figure 3). The antibodies were coated overnight at room temperature in PBS at a concentration of 10 μg/ml.

Clones found to give a positive ELISA signal were screened by PCR (Gussow and Clackson, 1989) and 'fingerprinted' with the restriction enzyme *Bst*NI (Clackson *et al.*, 1991) as in Marks *et al.* (1991) to identify different clones. Examples of clones with different restriction patterns were selected and the heavy and light chains sequenced (Sanger *et al.*, 1977) using a Sequenase kit (USB) or using a Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems) and an Applied Biosystems 373A DNA sequencer.

Sequenced clones were further analysed using the program MacVector 3.5 (IBI Kodak, New Haven, CT). The VH genes were compared with the 83 germline gene segments present in the VH directory compiled by Tomlinson *et al.* (1992). VL genes were compared with 34 published kappa germline gene segments (Bentley and Rabbits, 1980, 1983; Jaenichen *et al.*, 1984; Pech *et al.*, 1984, 1985; Pech and Zachau, 1984; Klobeck *et al.*, 1985a,b; Stavnezer *et al.*, 1985; Chen *et al.*, 1986, 1987a,b; Lorenz *et al.*, 1988; Straubinger *et al.*, 1988a,b; Scott *et al.*, 1989, 1991; Pargent *et al.*, 1991; Lautner *et al.*, 1992) and 13 published lambda gene segments (Anderson *et al.*, 1984; Alexandre *et al.*, 1989; Brockly *et al.*, 1989; Siminovitch *et al.*, 1989; Bernard *et al.*, 1990; Frippiat *et al.*, 1990; Combriato and Klobeck, 1991; Marks *et al.*, 1991; Daley *et al.*, 1992; Winkler *et al.*, 1992). Regions of the V-genes encoded by PCR primers were not included in the analysis.

Characterization of selected scFv fragments

The following clones were chosen for large scale purification and further characterization: αFOG1-H6, αFOG1-A3, αTNF-E7 and αThy-29. Colonies of the non-suppressor *E.coli* strain HB2151 harbouring the appropriate phagemid were used to inoculate 2 l of 2×TY containing 100 μg/ml ampicillin and 0.1% glucose. The cultures were grown and induced (De Bellis and Schwartz, 1990) and the tagged scFv fragments purified using the mAb 9E10 as in Clackson *et al.* (1991).

The inhibition of ^{125}I-Fog-1 binding to human Rh D antigen by the affinity purified scFv fragments αFOG1-H6 and αFOG1-A3 was essentially as performed earlier (Gorick *et al.*, 1988) with the following modifications. 0.0148 μg of ^{125}I-FOG1 was pre-incubated with varying amounts of purified αFOG1-H6 or αFOG1-A3 scFv fragments (0–16 μg) at 37°C for 1.5 h, before adding 0.5 μl of R_1R_2 cells (or rr cells as control). The mixture was then incubated for a further 1.5 h at 37°C with constant mixing and finally cells separated from the supernatant. As a control, a titration was also performed with a purified scFv fragment directed against turkey egg white lysozyme (αTEL9) (Marks *et al.*, 1991).

Kinetic measurements were made using surface plasmon resonance (BIAcore, Pharmacia Biosensor AB) (Jönsson *et al.*, 1991; Jönsson and Malmqvist, 1992). In order to separate monomeric and multimeric species, the purified scFv fragments were concentrated by ultrafiltration and then fractionated on a calibrated Superdex 75 FPLC column (Pharmacia) in PBS, 0.2 mM EDTA. Gel filtration was monitored both by the absorbance at 280 nm and on-line to BIAcore with immobilized antigen on the sensor chip (Johnsson *et al.*, 1991).

Kinetic experiments were performed in two different configurations. First, to analyse the binding of soluble scFv, the different antigens were covalently immobilized on the sensor chip (in the case of mAb Fog-1, the antibody was also immobilized via a mouse anti-human kappa light chain mAb using a sensor chip coated with rabbit anti-mouse IgG1). Secondly, to analyse the binding of the soluble mAb FOG-1, the αFOG1-H6 scFv was immobilized on the chip surface.

The antigens were coupled to the CM5 sensor chip through their amine groups using the Amine Coupling Kit (Pharmacia Biosensor AB) (Johnsson *et al.*, 1991). The antigens were diluted in 10 mM acetate buffer (pH 5.0) to ~25 μg/ml and 3805 resonance units (RU) of TNF, 6249 RU of human thyroglobulin and 5279 RU of FOG1 were immobilized. For the biospecific presentation of Fog-1, affinity purified rabbit anti-mouse IgG1 (Pharmacia Biosensor AB) was coupled to the surface followed by a mouse mAb anti-human kappa (2300 RU) and then Fog-1 (2050 RU). As binding of the rabbit anti-mouse IgG1 to the mouse mAb was reversible by 10 mM HCl the complex was rebuilt for each analytical cycle. ScFv anti-Fog-1 was coupled to the CM5 surface to 1538 RU. All determinations were performed at 25°C in PBS, 0.2 mM EDTA, 0.05% BIAcore surfactant P20 with a constant flow rate of 10 μl/min and an injected vol sample of 35 μl. It was not necessary to regenerate the antigen as the scFv fragments rapidly dissociate, with the exception of the biospecific presentation of antigen via rabbit anti-mouse IgG1 which was regenerated with 10 mM HCl for 3 min.

Analyses of scFv monomer were performed in the concentration range 100–500 nM and dimers in the range 40–200 nM except for the biospecifically presented Fog-1 where the concentration of dimeric scFv was 0.25–1.26 μM. Fog-1 was analysed on the αFOG1-H6 scFv surface in the concentration range 10–200 nM. All concentrations were calculated from U.V. absorption at 280 nm [assuming that 0.7 mg/ml scFv gives an A_{280} = 1 (Mach *et al.*, 1992) and that M_r of a scFv monomer is 30 kDa and of a dimer is 60 kDa]. No correction was made for the fraction of active protein, and therefore the on-rates are an underestimate. The kinetic evaluation of data was performed according to Karlsson *et al.* (1991) and evaluated on the program Origin 1.1 (Microcal inc., Northampton, MA, USA).

Acknowledgements

We thank M.R.Price for providing CEA, MUC1 peptide coupled to Sepharose 4B and PEM, A.Nissim, C.Chothia and S.Williams for help with sequence analysis, and Peptide Technology Ltd (Sydney, Australia) for a gift of recombinant TNFα. Recombinant soluble CD4 was provided by the MRC AIDS Reagent Project. A.D.G. and M.J.E. were supported by the Cancer Research Campaign, M.M. by Pharmacia Biosensor AB, J.D.M. by the MRC AIDS Directed Programme and the MRC, J.M.B., B.D.G. and N.C.H.-J. by a grant from the International Blood Group Reference Laboratory, J.McC. by Cambridge Antibody Technology Ltd., M.B. by the AIDS Program of the Federal Ministry for Research and Technology Germany, K.P.H. by ETH Zurich, and H.R.H. by the D.Collen Research Foundation, Leuven and the European Molecular Biology Organization.

References

Adelstein,S., Pritchard,B.H., Anderson,T.A., Crosbie,J., Gammon,G., Loblay,R.H., Basten,A. and Goodnow,C.C. (1991) *Science*, **251**, 1223–1225.

Alexandre,D., Chuchana,P., Brockly,F., Blancher,A., Lefranc,G. and Lefranc,M.P. (1989) *Nucleic Acids Res.*, **17**, 3975.

Anderson,M.L., Szajnert,M.F., Kaplan,J.C., McColl,L. and Young,B.D. (1984) *Nucleic Acids Res.*, **12**, 6647–6661.

Avrameas,S. (1991) *Immunol. Today*, **12**, 154–159.

Basten,A., Brink,R., Peake,P., Adams,E., Crosbie,J., Hartley,S. and Goodnow,C.C. (1991) *Immunol. Rev.*, **122**, 5–19.

Bendtzen,K., Svenson,M., Jnsson,V. and Hippe,E. (1990) *Immunol. Today*, **11**, 167–169.

Bentley,D.L. and Rabbits,T.H. (1980) *Nature*, **288**, 730–733.

Bentley,D.L. and Rabbitts,T.H. (1983) *Cell*, **32**, 181–189.

Berek,C. and Milstein,C. (1987) *Immunol. Rev.*, **96**, 23–41.

Bernard,F., Chuchana,P., Frippiat,J.P., Buluwela,L. and Lefranc,M.P. (1990) *Nucleic Acids Res.*, **18**, 7139.

Bird,R.E. *et al.* (1988) *Science*, **242**, 423–426.

Bona,C.A. (1988) *Annu. Rev. Immunol.*, **6**, 327–358.

Bouanani,M., Bataille,R., Piechaczyk,M., Salhi,S.L., Pau,B. and Bastide,M. (1991) *Arthritis Rheumatol.*, **34**, 1585–1593.

Bretscher,P. and Cohn,M. (1970) *Science*, **169**, 1042–1049.

Brockly,F., Alexandre,D., Chuchana,P., Huck,S., Lefranc,G. and Lefranc,M.P. (1989) *Nucleic Acids Res.*, **17**, 3976.

Burnet,F.M. (1959) *The Clonal Selection Theory of Acquired Immunity*. Vanderbilt University Press, Nashville.

Bye,J.M., Carter,C., Cui,Y., Gorick,B.D., Songsivilai,S., Winter,G., Hughes-Jones,N.C. and Marks,J.D. (1993) *J. Clin. Invest.*, in press.

Casali,P. and Notkins,A.L. (1989) *Annu. Rev. Immunol.*, **7**, 513–535.

Chen,P.P., Albrandt,K., Orida,N.K., Radoux,V., Chen,E.Y., Schrantz,R., Liu,F.T. and Carson,D.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8318–8322.

Chen,P.P., Albrandt,K., Kipps,T.J., Radoux,V., Liu,F.-T. and Carson,D.A. (1987a) *J. Immunol.*, **139**, 1727–1733.

Chen,P.P., Robbins,D.L., Jirik,F.R., Kipps,T.J. and Carson,D.A. (1987b) *J. Exp. Med.*, **166**, 1900–1905.

Clackson,T., Hoogenboom,H.R., Griffiths,A.D. and Winter,G. (1991) *Nature*, **352**, 624–628.

Cohen,I.R. and Cooke,A. (1986) *Immunol. Today*, **7**, 363–364.

Combriato,G. and Klobeck,H.G. (1991) *Eur. J. Immunol.*, **21**, 1513–1522.

Daley,D.M., Olee,T., Peng,H.-Q., Soto-Gil,R.W., Chen,P.P. and Siminovitch,K.A. (1992) *Mol. Immunol.*, **29**, 1031–1042.

De Bellis,D. and Schwartz,I. (1990) *Nucleic Acids Res.*, **18**, 1311.

Erikson,J., Radic,M.Z., Camper,S.A., Hardy,R.R., Carmack,C. and Weigert,M. (1991) *Nature*, **349**, 331–334.

Foote,J. and Milstein,C. (1991) *Nature*, **352**, 530–532.

Foote,J. and Winter,G. (1992) *J. Mol. Biol.*, **224**, 487–499.

Trippier,J.F., Chuchana,P., Bernard,P., Bonnefoi,L., Lefranc,G. and Lefranc,M.P. (1990) *Nucleic Acids Res.*, **18**, 7134.
Gendler,S., Taylor-Papadimitrou,J., Duhig,T., Rothbard,J. and Burchell,J. (1988) *J. Biol. Chem.*, **263**, 12820–12823.
Gherardi,E. and Milstein,C. (1992) *Nature*, **357**, 201–202.
Gibson,T.J. (1984) PhD thesis, University of Cambridge, UK.
Glockshuber,R., Malia,M., Pfitzinger,I. and Pluckthun,A. (1990) *Biochemistry*, **29**, 1362–1367.
Gorick,B.D., Thompson,K.M., Melamed,M.D. and Hughes,J.N. (1988) *Vox. Sang.*, **55**, 165–170.
Gram,H., Marconi,L., Barbas,C.F., Collet,T.A., Lerner,R.A. and Kang,A.S. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 3576–3580.
Griffiths,G.M., Berek,C., Kaartinen,M. and Milstein,C. (1984) *Nature*, **312**, 271–275.
Gum,J.R., Hicks,J.W., Swallow,D.M., Lagace,R.L., Byrd,J.C., Lamport,D.T.A., Siddiki,B. and Kim,Y.S. (1990) *Biochem. Biophys Res. Commun.*, **171**, 407–415.
Gussow,D. and Clackson,T. (1989) *Nucleic Acids Res.*, **17**, 4000.
Hale,G., Dyer,M.J., Clark,M.R., Phillips,J.M., Marcus,R., Riechmann,L., Winter,G. and Waldmann,H. (1988) *Lancet*, **2**, 1394–1399.
Hartley,S.B., Crosbie,J., Brink,R., Kantor,A.B., Basten,A. and Goodnow,C.C. (1991) *Nature*, **353**, 765–769.
Hawkins,R.E. and Winter,G. (1992) *Eur. J. Immunol.*, **22**, 867–870.
Holmberg,D. and Coutinho,A. (1985) *Immunol. Today*, **6**, 356–357.
Hoogenboom,H.R., Griffiths,A.D., Johnson,K.S., Chiswell,D.J., Hudson,P. and Winter,G. (1991) *Nucleic Acids Res.*, **19**, 4133–4137.
Hoogenboom,H.R., Marks,J.D., Griffiths,A.D. and Winter,G. (1992) *Immunol. Rev.*, in press.
Huse,W.D., Sastry,L., Iverson,S.A., Kang,A.S., Alting,M.M., Burton,D.R., Benkovic,S.J. and Lerner,R.A. (1989) *Science*, **246**, 1275–1281.
Huston,J.S. *et al.* (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5879–5883.
Jaenichen,H.R., Pech,M., Lindenmaier,W., Wildgruber,N. and Zachau,H.G. (1984) *Nucleic Acids Res.*, **12**, 5249–5263.
James,K. and Bell,G.T. (1987) *J. Immunol. Methods*, **100**, 5–40.
Johnsson,B., Löfås,S. and Lindqvist,G. (1991) *Anal. Biochem.*, **198**, 268–277.
Jones,D.H., McMillan,A.J., Fersht,A.R. and Winter,G. (1985) *Biochemistry*, **24**, 5852–5857.
Jönsson,U. and Malmqvist,M. (1992) In Turner,A. (ed.), *Real Time Biospecific Interaction*. JAI Press Ltd, San Diego, Vol 2, pp. 291–336.
Jönsson,U. *et al.* (1991) *BioTechniques*, **11**, 620–627.
Karlsson,R., Michaelsson,A. and Mattsson,L. (1991) *J. Immunol. Methods*, **145**, 229–240.
Klobeck,H.G., Bornkamm,G.W., Combriato,G., Mocikat,R., Pohlenz,H.D. and Zachau,H.G. (1985a) *Nucleic Acids Res.*, **13**, 6515–6529.
Klobeck,H.G., Meindl,A., Combriato,G., Solomon,A. and Zachau,H.G. (1985b) *Nucleic Acids Res.*, **13**, 6499–6513.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Lautner,R.A., Huber,C., Meindl,A., Pargent,W., Schable,K.F., Thiebe,R., Zocher,I. and Zachau,H.G. (1992) *Eur. J. Immunol.*, **22**, 1023–1029.
Lorenz,W., Schable,K.F., Thiebe,R., Stavnezer,J. and Zachau,H.G. (1988) *Mol. Immunol.*, **25**, 479–484.
Mach,H., Middaugh,C.R. and Lewis,R.V. (1992) *Anal. Biochem.*, **200**, 74–80.
Mach,J.-P., Forni,M., Ritschard,J., Buchegger,F., Carrel,S., Widgren,S., Donath,A. and Alberto,P. (1980) *Oncodevel. Biol. Med.*, **1**, 49–69.
Malthiéry,Y. and Lissitzky,S. (1987) *Eur. J. Biochem.*, **165**, 491–498.
Marks,J.D., Hoogenboom,H.R., Bonnert,T.P., McCafferty,J., Griffiths,A.D. and Winter,G. (1991) *J. Mol. Biol.*, **222**, 581–597.
Marks,J.D., Hoogenboom,H.R., Griffiths,A.D. and Winter,G. (1992a) *J. Biol. Chem.* **267**, 16007–16010.
Marks,J.D., Griffiths,A.D., Malmqvist,M., Clackson,T., Bye,J.M. and Winter,G. (1992b) *Bio/Technology*, **10**, 779–783.
Mason,D.W. and Williams,A.F. (1986) *Kinetics of Antibody Reactions and the Analysis of Cell Surface Antigens*. Blackwell Scientific, Oxford.
McCafferty,J., Griffiths,A.D., Winter,G. and Chiswell,D.J. (1990) *Nature*, **348**, 552–554.
Melamed,M.D., Thompson,K.M., Gibson,T. and Hughes-Jones,N.C. (1987) *J. Immunol. Methods*, **104**, 245–251.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Munro,S. and Pelham,H.R.B. (1986) *Cell*, **46**, 291–300.
Nakamura,M., Burastero,S.E., Ueki,Y., Larrick,J.W., Notkins,A.L. and Casali,P. (1988) *J. Immunol.*, **141**, 4165–4172.
Nemazee,D., Russell,D., Arnold,B., Haemmerling,G., Allison,J., Miller,J.F., Morahan,G. and Buerki,K. (1991) *Immunol. Rev.*, **122**, 117–132.
Nossal,G.J. (1987) *Int. Rev. Immunol.*, **2**, 321–338.
Nossal,G.J. (1989) *Science*, **245**, 147–153.
Pargent,W., Meindl,A., Thiebe,R., Mitzel,S. and Zachau,H.G. (1991) *Eur. J. Immunol.*, **21**, 1821–1827.
Pech,M. and Zachau,H.G. (1984) *Nucleic Acids Res.*, **12**, 9229–9236.
Pech,M., Jaenichen,H.R., Pohlenz,H.D., Neumaier,P.S., Klobeck,H.G. and Zachau,H.G. (1984) *J. Mol. Biol.*, **176**, 189–204.
Pech,M., Smola,H., Pohlenz,H.D., Straubinger,B., Gerl,R. and Zachau,H.G. (1985) *J. Mol. Biol.*, **183**, 291–299.
Pecht,I. (1982) In Sela,M. (ed.), *Dynamic Aspects of Antibody Function*. Academic Press, Inc., New York, Vol 6, pp. 1–68.
Perelson,A.S. and Oster,G.F. (1979) *J. Theor. Biol.*, **81**, 645–670.
Persson,M.A.A., Caothien,R.H. and Burton,D.R. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 2432–2436.
Price,M.R., Hudecz,F., O'Sullivan,C., Baldwin,R.W., Edwards,P.M. and Tendler,S.J.B. (1990) *Mol. Immunol.*, **27**, 795–802.
Rossi,F., Guilbert,B., Tonnelle,C., Ternynck,T., Fumoux,F., Avrameas,S. and Kazatchkine,M.D. (1990) *Eur. J. Immunol.*, **20**, 2089–2094.
Ruf,J., Carayon,P. and Lissitzky,S. (1985) *Eur. J. Immunol.*, **15**, 268–272.
Russell,D.M., Dembic,Z., Morahan,G., Miller,J.F., Burki,K. and Nemazee,D. (1991) *Nature*, **354**, 308–311.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1990) *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Scott,M.G., Crimmins,D.L., McCourt,D.W., Zocher,I., Thiebe,R., Zachau,H.G. and Nahm,M.H. (1989) *J. Immunol.*, **143**, 4110–4116.
Scott,M.G., Crimmins,D.L., McCourt,D.W., Chung,G., Schable,K.F., Thiebe,R., Quenzel,E.M., Zachau,H.G. and Nahm,M.H. (1991) *J. Immunol.*, **147**, 4007–4013.
Siminovitch,K.A., Misener,V., Kwong,P.C., Song,Q.-L. and Chen,P.P. (1989) *J. Clin. Invest.*, **84**, 1675–1678.
Smith,G.P. (1985) *Science*, **228**, 1315–1317.
Smith,H.R. and Steinberg,A.D. (1983) *Annu. Rev. Immunol.*, **1**, 175–210.
Smith,T.W. and Skubitz,K.M. (1975) *Biochemistry*, **14**, 1496–1502.
Spooner,C.E., Markowitz,N.P. and Saravolatz,L.D. (1992) *Clin. Immunol. Immunopathol.*, **62**, S11–17.
Stavnezer,J., Kekish,O., Batter,D., Grenier,J., Balazs,I., Henderson,E. and Zegers,B.J. (1985) *Nucleic Acids Res.*, **13**, 3495–514.
Straubinger,B., Huber,E., Lorenz,W., Osterholzer,E., Pargent,W., Pech,M., Pohlenz,H.D., Zimmer,F.J. and Zachau,H.G. (1988a) *J. Mol. Biol.*, **199**, 23–34.
Straubinger,B., Thiebe,R., Huber,C., Osterholzer,E. and Zachau,H.G. (1988b) *Biol. Chem. Hoppe-Seyler*, **369**, 601–607.
Thompson,K.M., Hough,D.W., Maddison,P.J., Melamed,M.D. and Hughes-Jones, N.C. (1986) *J. Immunol. Methods*, **94**, 7–12.
Tomer,Y. and Schoenfeld,Y. (1988) *Immunol. Invest.*, **17**, 389–424.
Tomlinson,I.M., Walter,G., Marks,J.D., Llewelyn,M.B. and Winter,G. (1992) *J. Mol. Biol.*, **227**, 776–798.
Ward,E.S., Gussow,D., Griffiths,A.D., Jones,P.T. and Welch,M.J., Fong,S., Vaughan,J. and Carson,D. (1983) *Clin. Exp. Immunol.*, **51**, 299–304.
Winkler,T.H., Fehr,H. and Kalden,J.R. (1992) *Eur. J. Immunol.*, **22**, 1719–1728.
Winter,G. (1989) *Nature*, **341**, 544–546.
Winter,G. and Milstein,C. (1991) *Nature*, **349**, 293–299.

Received on September 29, 1992; revised on November 11, 1992

Note added in proof

Nucleotide sequences of all the V-genes whose deduced protein sequences are given in Table II have been submitted to the EMBL Data Library and assigned the following accession numbers: FOG1VHA3, Z18822; FOG1VLA3, Z18823; FOG1VHH6, Z18824; FOG1VLH6, Z18825; FOG1VHA4, Z18826; FOG1VLA4, Z18827; FOG1VHG8, Z18828; FOG1VLG8, Z18829; THYVH23, Z18830; THYVL23, Z18831; THYVH29, Z18832; THYVL29, Z18833; THYVH32, Z18834; THYVL32. Z18835; THYVH33, Z18836; THYVL33, Z18837; TNFVLA1, Z18838; TNFVHE1, Z18839; TNFVLE1, Z18840; TNFVHE7, Z18841; TNFVLE7, Z18842; TNFVHH9, Z18843; TNFVLH9, Z18844; CEAVL8A, Z18845; MUC1VH1, Z18846; MUC1VL1, Z18847; CD4VH74, Z18848; CD4VL74, Z18849; TNFVHA1, Z18850; CEAVH8A, Z18851.

66. T. A. Stewart, P. K. Pattengale, P. Leder, *Cell* **38**, 627 (1984).
67. C. Schoenenberger *et al.*, *EMBO J.* **7**, 169 (1988).
68. U. Ruther, C. Garber, D. Komitowski, R. Muller, E. F. Wagner, *Nature* **325**, 412 (1987); U. Ruther, D. Komitowski, F. R. Schubert, E. F. Wagner, *Oncogene* **4**, 861 (1989).
69. R. L. Brinster *et al.*, *Cell* **37**, 367 (1984); R. D. Palmiter, H. Y. Chen, A. Messing, R. L. Brinster, *Nature* **316**, 457 (1985); T. A. Van Dyke *et al.*, *J. Virol.* **61**, 2029 (1987).
70. D. M. Ornitz *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **50**, 399 (1985).
71. S. Efrat and D. Hanahan, *Mol. Cell. Biol.* **7**, 192 (1987); S. Efrat, S. Baekkeskov, D. Lane, D. Hanahan, *EMBO J.* **6**, 2699 (1987).
72. K. A. Mahon *et al.*, *Science* **235**, 1622 (1987).
73. L. J. Field, *ibid.* **239**, 1029 (1988).
74. E. Sinn *et al.*, *Cell* **49**, 465 (1987).
75. C. J. Quaife *et al.*, *ibid.* **48**, 1023 (1987).
76. A. S. Tsukamoto *et al.*, *ibid.* **55**, 619 (1988).
77. W. J. Muller *et al.*, *ibid.* **54**, 105 (1988).
78. L. Bouchard, L. Lamarre, P. J. Tremblay, P. Jolicoeur, *ibid.* **57**, 931 (1989).
79. V. L. Bautch, S. Toda, J. A. Hassell, D. Hanahan, *ibid.* **51**, 529 (1987).
80. R. L. Williams, S. A. Courtneidge, E. F. Wagner, *ibid.* **52**, 121 (1988); R. L. Williams *et al.*, *ibid.* **57**, 1053 (1989); E. F. Wagner, R. L. Williams, U. Ruther, in *Cell to Cell Signals in Mammalian Development*, S. W. deLaat *et al.*, Eds. (Springer-Verlag, Berlin, 1989), pp. 301–310.
81. J. Vogel, S. H. Hinrichs, R. K. Reynolds, P. A. Luciw, G. Jay, *Nature* **335**, 606 (1988).
82. S. H. Hinrichs, M. Nerenberg, R. K. Reynolds, G. Khoury, G. Jay, *Science* **237**, 1340 (1987); M. Nerenberg, S. H. Hinrichs, R. K. Reynolds, G. Khoury, G. Jay, *ibid.*, p. 1324.
83. M. Lacey, S. Alpert, D. Hanahan, *Nature* **322**, 609 (1986); M. Sippola-Thiele, D. Hanahan, P. M. Howley, *Mol. Cell. Biol.* **9**, 925 (1989).
84. I thank I. Herskowitz, T. DeFranco, and M. Kirschner for many thoughtful comments on the manuscript, and L. Spector for manuscript preparation.

Research Article

Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda

WILLIAM D. HUSE, LAKSHMI SASTRY, SHEILA A. IVERSON, ANGRAY S. KANG, MICHELLE ALTING-MEES, DENNIS R. BURTON, STEPHEN J. BENKOVIC, RICHARD A. LERNER

A novel bacteriophage lambda vector system was used to express in *Escherichia coli* a combinatorial library of Fab fragments of the mouse antibody repertoire. The system allows rapid and easy identification of monoclonal Fab fragments in a form suitable for genetic manipulation. It was possible to generate, in 2 weeks, large numbers of monoclonal Fab fragments against a transition state analog hapten. The methods described may supersede present-day hybridoma technology and facilitate the production of catalytic and other antibodies.

MONOCLONAL ANTIBODIES HAVE BEEN GENERATED THAT catalyze chemical transformations ranging from simple acyl transfer reactions to the energetically demanding hydrolysis of the peptide bond in the presence of metal cofactors (*1, 2–11*). Initially, it was widely held that antibodies would be most useful for catalysis where their predominant role was to overcome entropic barriers that occur along the reaction pathway. The basis of this hypothesis was that the chance occurrence of amino acid side chains capable of acid base catalysis in proximity to the reaction center was unlikely. However, for some reactions, study of the *p*H rate profile has revealed the participation of monobasic residues. Other studies have focused on placing appropriate charges on the antigen to induce specific binding interactions by complementary charged amino acid side chains on the antibody (*9, 12, 13*). Such functionalities might participate as a general acid, base, or nucleophile in the reaction under study.

Apart from the validity of the design of the mechanism based antigen, the probability of finding antibodies where particular amino acid side chains participate in catalysis also depends on the number of different antibodies assayed. Because current methods of generating monoclonal antibodies do not provide for an adequate survey of the available repertoire, we have been devising methods to clone the antibody repertoire in *Escherichia coli* and have described the preparation of a highly diverse immunoglobulin gene library (*14*). Given the difficulty of expressing both heavy and light chains together, we initially considered the construction and expression of libraries restricted to fragments of the variable region of the immunoglobulin (Ig) heavy chain V_H (*14*). In fact, a recent report describes the construction of a plasmid expression library in *E. coli* in which V_H fragments with affinity for keyhole limpet hemocyanin (KLH) and lysozyme have been isolated (*15*). However, the use of isolated V_H fragments as antibody mimics may be limited because (i) the available crystal structures of antibody-antigen complexes show considerable contact between antigen and V_L (light chain

W. D. Huse, L. Sastry, S. A. Iverson, and R. A. Lerner are with the Departments of Molecular Biology and Chemistry, Research Institute of Scripps Clinic, La Jolla, CA 92037. D. R. Burton and A. S. Kang are at the Departments of Molecular Biology and Chemistry, Research Institute of Scripps Clinic, La Jolla, CA 92037 and at the Krebs Institute, Department of Molecular Biology and Biotechnology, The University, Sheffield, United Kingdom. M. Alting-Mees is at Stratagene Inc., La Jolla, CA 92037. S. J. Benkovic is at the Department of Chemistry, Pennsylvania State University, University Park, PA 16802 and at the Departments of Molecular Biology and Chemistry, Research Institute of Scripps Clinic, La Jolla, CA 92037

variable) domain as well as V_H (*16*). More explicitly, in the case of a series of antibodies to dextran, the V_L domain provides contacts critical to antigen binding (*17*). Thus, it is unlikely that the affinity of isolated V_H fragments will generally match that of intact antibodies. (ii) The absence of the V_L domain leaves a large hydrophobic patch on one face of the V_H fragment, which will almost certainly lead to increased nonspecificity relative to whole antibodies (*15*). In contrast, Fab fragments (antigen binding) have been studied for more than 30 years. They behave as whole antibodies in terms of antigen recognition, and their affinity and specificity are well defined. Furthermore, for Fab, the combinatorial properties of heavy and light chains serve as an important source of diversity.

In that individual Fab molecules can be expressed and assembled in *E. coli* (*18*), the route to mimicking the diversity of the antibody system in vitro should lie in solving the problem of expressing the repertoires of heavy and light chains in combination. Accordingly, we used a novel system to enable the construction of bacteriophage lambda (λ) libraries expressing a population of functional antibody fragments (Fab's) with a potential diversity equal to or exceeding that of the parent animal.

Criteria for vector construction. To obtain a vector system for generating the largest number of Fab fragments that could be screened directly, we constructed the expression libraries in bacteriophage λ for the following reasons. First, in vitro packaging of phage DNA is the most efficient method of reintroducing DNA into host cells. Second, it is possible to detect protein expression at the level of single-phage plaques. Finally, in our experience, screens of phage libraries diminish the usual difficulties with nonspecific binding. The alternative, plasmid cloning vectors are only advantageous in the analysis of clones after they have been identified. This advantage is not lost in our system because we use λzap II and are able to excise a plasmid (*19*) containing the heavy chain, light chain, or Fab expressing inserts.

The vectors for expression of V_H, V_L, Fv (fragment of the variable region), and Fab sequences are diagrammed in Figs. 1 and 2. They were constructed by a modification of λzap II (*19*) in which we inserted synthetic oligonucleotides into the multiple cloning site. The vectors were designed to be antisymmetric with respect to the Not I and Eco RI restriction sites that flank the cloning and expression sequences. This antisymmetry in the placement of restriction sites in a linear vector such as bacteriophage allows a library expressing light chains to be combined with one expressing heavy chains in order to construct combinatorial Fab expression libraries. The vector λLc1 is designed to serve as a cloning vector for light

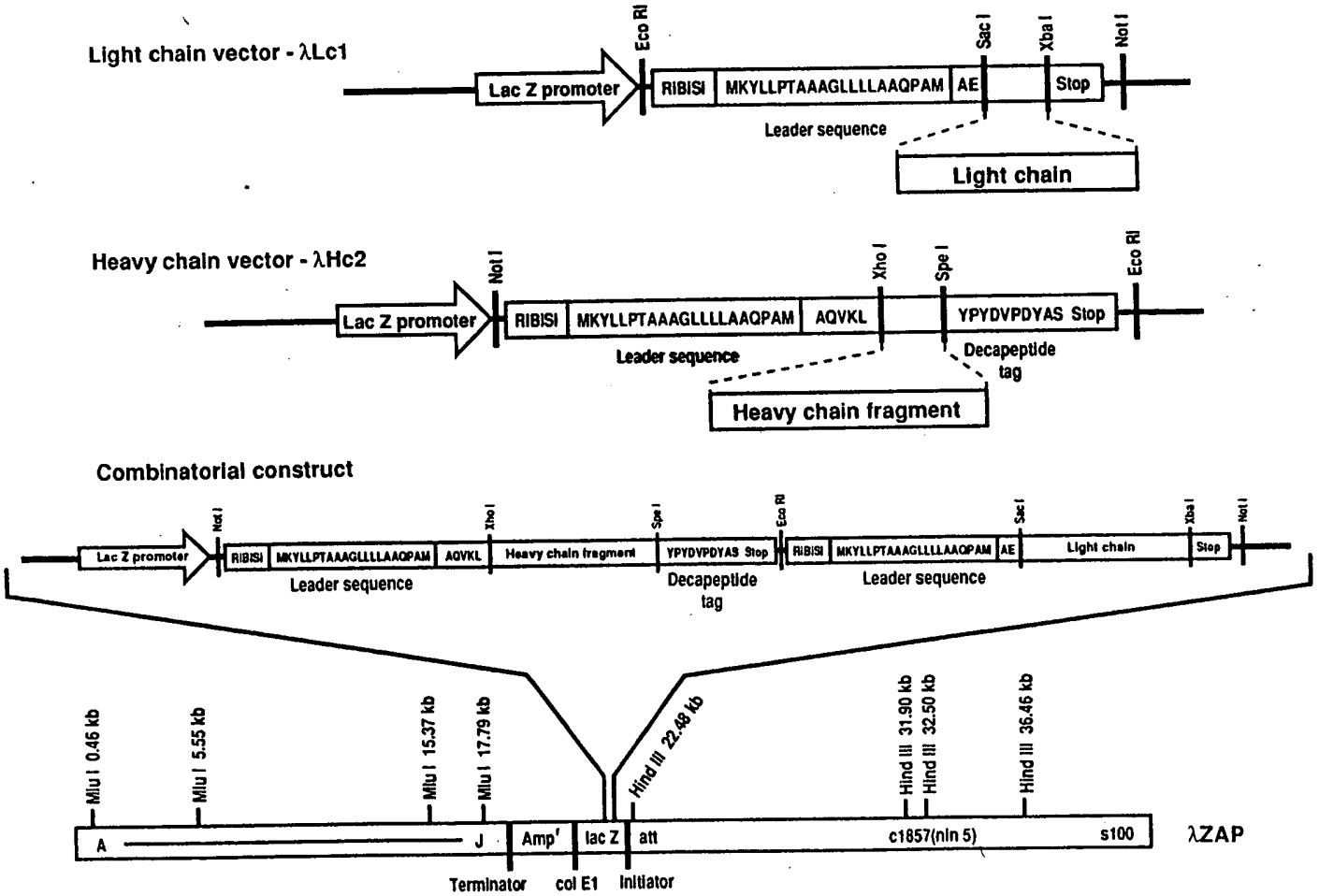


Fig. 1. Combinatorial bacteriophage λ vector system for expression of Fab antibody fragments. The λLc1 vector was constructed for the cloning of PCR amplified products of mRNA that code for light chain protein by inserting the nucleotide sequence shown depicted in Fig. 2A into the Sac I and Xho I sites of λzap II. The sequence was constructed from overlapping synthetic oligonucleotides varying in length from 25 to 50 nucleotides. The λHc2 vector was constructed for cloning PCR amplified products coding for heavy chain Fd sequences by inserting the nucleotide sequence (Fig. 2B) into the Not I and Xho I sites of λzap II. As with the light chain vector, the inserted sequence was constructed from overlapping synthetic oligonucleotides. The combinatorial constructs that can express Fab fragments are generated by cutting DNA isolated from light and heavy chain libraries at the antisymmetric Eco RI site of each vector, followed by re-ligation of the resulting arms. This generates constructs having random combination of light and heavy chains which can be expressed, upon induction with IPTG, from a dicistronic mRNA from the lac Z promoter.

chain fragments, and λHc2 is designed to serve as a cloning vector for heavy chain sequences in the initial step of library construction. These vectors are engineered to efficiently clone the products of PCR amplification with specific restriction sites incorporated at each end (*14, 15*). The sequence of the oligonucleotides used to construct these vectors include elements for construction, expression, and secretion of Fab fragments. These oligonucleotides introduce the antisymmetric Not I and Eco RI restriction sites; a leader peptide for the bacterial *pel* B gene, which has previously been successfully used in *E. coli* to secrete Fab fragments (*18*); a ribosome binding site at the optimal distance for expression of the cloned sequence; cloning sites for either the light or heavy chain PCR product; and, in λHc2, a decapeptide tag at the carboxyl terminus of the expressed heavy chain protein fragment. The sequence of the decapeptide tag was useful because of the availability of monoclonal antibodies to this peptide that were used for immunoaffinity purification of fusion proteins (*20*). The restriction endonuclease recognition sites included in the vectors were Sac I and Xba I in λLc1, and Xho I and Spe I in λHc2. The vectors were characterized by restriction digest analysis and DNA sequencing.

Choice of antigen and amplification of antibody fragments. We constructed the initial Fab expression library from mRNA isolated from a mouse that had been immunized with the KLH-coupled *p*-nitrophenyl phosphonamidate antigen 1 (NPN) (Fig. 3). This antigen was shown by Janda and co-workers (*7*) to be an effective one for the generation of catalytic antibodies. Also, the antibodies for the NPN reaction have been identified and therefore facilitate the implementation of assay systems. Finally, successful generation of catalytic antibodies generally requires binding to relatively small organic haptens, and it was necessary to test the suitability of our system for such molecules.

The PCR amplification of messenger RNA (mRNA) isolated from spleen cells or hybridomas with oligonucleotides that incorporate restriction sites into the ends of the amplified product can be used to clone and express heavy chain sequences (*14, 15*). This work is now extended to include the amplification of the Fd (V_H − C_H1)

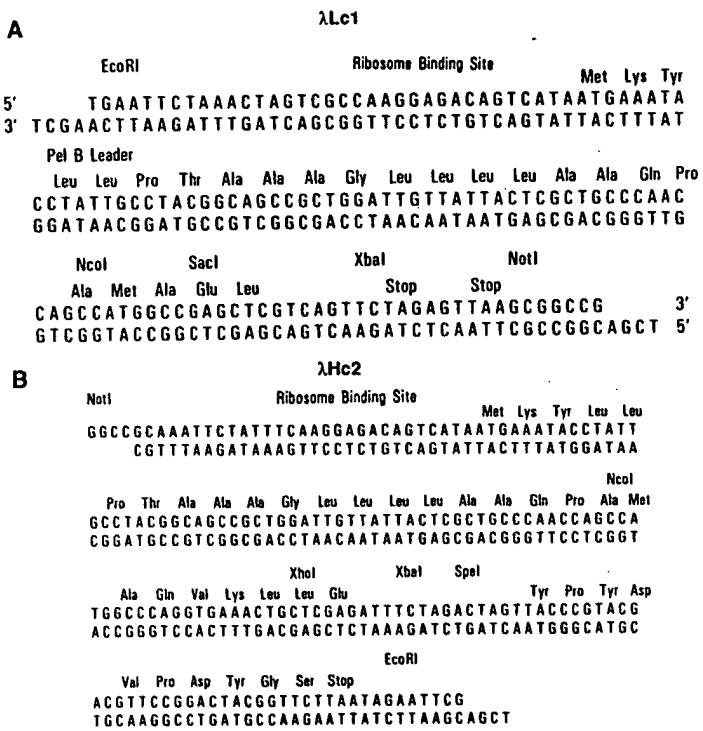


Fig. 2. (**A**) The nucleotide sequence inserted into λzap II to construct λLc1. (**B**) The nucleotide sequence inserted into λzap II to construct λHc2.

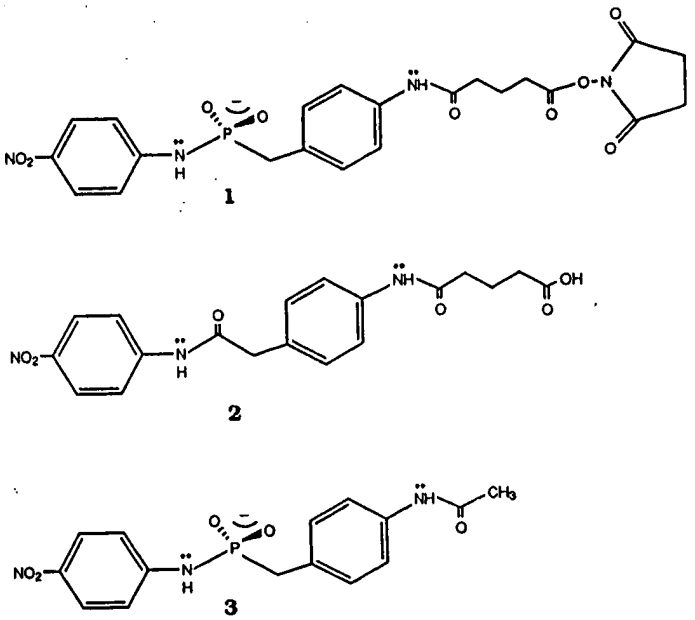


Fig. 3. The transition state analog **1**, which induces antibodies for hydrolyzing carboxamide substrate **2**. Compound **1** containing a glutaryl spacer and an *N*-hydroxysuccinimide–linker appendage is the form used to couple the hapten **1** to protein carriers KLH and BSA, while **3** is the inhibitor. The phosphonamidate functionality is a mimic of the stereoelectronic features of the transition state for hydrolysis of the amide bond.

and κ chain sequences (Fig. 4) from mouse spleen cells. The oligonucleotide primers used for these amplifications (Tables 1 and 2) are analogous to those that have been successfully used for amplification of V_H sequences (*14*). The set of 5′ primers for heavy chain amplification was identical to those used to amplify V_H, and those for light chain amplification were chosen similarly (*14, 21*). The 3′ primers of heavy (IgG1) and light (κ) chain sequences included the cysteines involved in disulfide bond formation between heavy and light chains. At this stage no primer was constructed to amplify light (λ) chains since they constitute only a small fraction of murine antibodies (*22*). Restriction endonuclease recognition sequences were incorporated into the primers to allow for the cloning of the amplified fragment into a λ phage vector in a predetermined reading frame for expression.

Library construction. We constructed a combinatorial library in two steps. In the first step, separate heavy and light chain libraries were constructed in λHc2 and λLc1, respectively (Fig. 1). In the second step, these two libraries were combined at the antisymmetric Eco RI sites present in each vector. This resulted in a library of clones each of which potentially coexpresses a heavy and a light chain. The actual combinations are random and do not necessarily reflect the combinations present in the B cell population in the parent animal. The λHc2 expression vector has been used to create a library of heavy chain sequences from DNA obtained by PCR amplification of mRNA isolated from the spleen of a 129 G_{IX}^+ mouse previously immunized with NPN conjugated to KLH. This primary library contains 1.3×10^6 plaque-forming units (pfu) and has been screened for the expression of the decapeptide tag to determine the percentage of clones expressing Fd sequences. The sequence for this peptide is only in frame for expression after the genes for an Fd (or V_H) fragment have been cloned into the vector. At least 80 percent of the clones in the library express Fd fragments when assayed by immunodetection of the decapeptide tag.

The light chain library was constructed in the same way as the heavy chain and shown to contain 2.5×10^6 members. Plaque screening, with an antibody to κ chain, indicated that 60 percent of

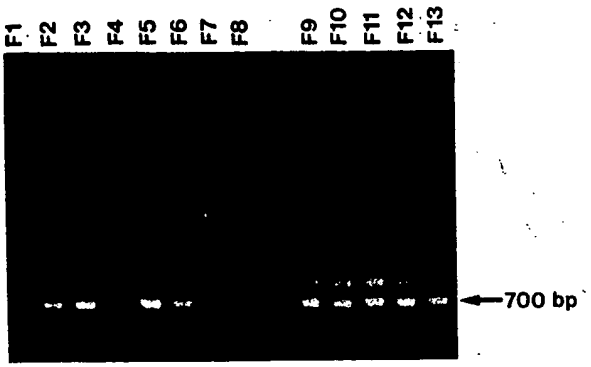


Fig. 4. PCR amplification of Fd and κ regions from the spleen mRNA of a mouse immunized with NPN. Amplification was performed as described (*14*) with RNA-cDNA hybrids obtained by the reverse transcription of the mRNA with primer specific for amplification of heavy chain sequences (12, Table 1) or light chain sequences (9, Table 2). Lanes F1 to F8 represent the product of heavy chain amplification reactions with one of each of the eight 5′ primers (primers 2 to 9, Table 1) and the 3′ primer (primer 12, Table 1). Light chain (κ) amplifications with the 5′ primers (primers 3 to 7, Table 2) and the appropriate 3′ primer (9, Table 2) are shown in lanes F9 through F13). A band of 700 base pairs is seen in all lanes indicating the successful amplification of Fd and κ regions.

HEAVY CHAIN PRIMERS
1) 5′-AGGT(C/G)(C/A)A(G/A)CT(G/T)CTCGAGTC(T/A)GG-3′
2) 5′-AGGTCCAGCTGCTCGAGTCTGG-3′
3) 5′-AGGTCCAGCTGCTCGAGTCAGG-3′
4) 5′-AGGTCCAGCTTCTCGAGTCTGG-3′
5) 5′-AGGTCCAGCTTCTCGAGTCAGG-3′
6) 5′-AGGTCCAACTGCTCGAGTCTGG-3′
7) 5′-AGGTCCAACTGCTCGAGTCAGG-3′
8) 5′-AGGTCCAACTTCTCGAGTCTGG-3′
9) 5′-AGGTCCAACTTCTCGAGTCAGG-3′
10) 5′-AGGTIIAICTICTCGAGTC(T/A)GG-3′
11) 5′-CTATTAACTAGTAACGGTAACAGTGGTGCCTTGCCCCA-3′
12) 5′-AGGCTTACTAGTACAATCCCTGGGCACAAT-3′

Table 1. Primers used for amplification of heavy chain Fd fragments for construction of Fab's. Amplification was performed in eight separate reactions, each containing one of the 5′ primers (primers 2 to 9) and one of the 3′ primers (primer 12). The remaining 5′ primers that were used for amplification in a single reaction are either a degenerate primer (primer 1) or a primer that incorporates inosine at four degenerate positions (primer 10). The remaining 3′ primer (primer 11) has been used to construct Fv fragments. The underlined portion of the 5′ primers incorporates an Xho I site and that of the 3′ primer on Spe I restriction site.

Table 2. Primers used for amplification of κ light chain sequences for construction of Fab's. Amplification was performed in five separate reactions, each containing one of the 5′ primers (primers 3 to 7) and one of the 3′ primers (primer 9). The remaining 3′ primer (primer 8) has been used to construct Fv fragments. The underlined portion of the 5′ primers incorporate a Sac I restriction site and that of the 3′ primers an Xba I restriction site.

LIGHT CHAIN PRIMERS
1) 5′-CCAGTTCCGAGCTCGTTGTGACTCAGGAATCT-3′
2) 5′-CCAGTTCCGAGCTCGTGTTGACGCAGCCGCCC-3′
3) 5′-CCAGTTCCGAGCTCGTGCTCACCCAGTCTCCA-3′
4) 5′-CCAGTTCCGAGCTCCAGATGACCCAGTCTCCA-3′
5) 5′-CCAGATGTGAGCTCGTGATGACCCAGACTCCA-3′
6) 5′-CCAGATGTGAGCTCGTCATGACCCAGTCTCCA-3′
7) 5′-CCAGTTCCGAGCTCGTGATGACACAGTCTCCA-3′
8) 5′-GCAGCATTCTAGAGTTTCAGCTCCAGCTTGCC-3′
9) 5′-GCGCCGTCTAGAATTAACACTCATTCCTGTTGAA-3′

the library contained expressed light chain inserts. This relatively small percentage of inserts probably resulted from incomplete dephosphorylation of the vector after cleavage with Sac I and Xba I.

Once obtained, the two libraries were used to construct a combinatorial library by crossing them at the Eco RI site as follows. DNA was first purified from each library. The light chain library was cleaved with Mlu I restriction endonuclease, the resulting 5′ ends were dephosphorylated, and the product was digested with Eco RI. This process cleaved the left arm of the vector into several pieces, but the right arm containing the light chain sequences remained intact. The DNA of heavy chain library was cleaved with Hind III, dephosphorylated, and then cleaved with Eco RI; this process destroyed the right arm, but the left arm containing the heavy chain sequences remained intact. The DNA's so prepared were then mixed and ligated. After ligation, only clones that resulted from combination of a right arm of light chain–containing clones and a left arm of heavy chain–containing clones reconstituted a viable phage. After ligation and packaging, 2.5×10^7 clones were obtained. This is the combinatorial Fab expression library that was screened to identify clones having affinity for NPN. For determining the frequency of the phage clones that coexpress the light and heavy chain fragments, we screened duplicate lifts of the combinatorial library for light and heavy chain expression. In our examination of approximately 500 recombinant phage, approximately 60 percent coexpressed light and heavy chain proteins.

Antigen binding. All three libraries, the light chain, the heavy chain, and Fab were screened to determine whether they contained recombinant phage that expressed antibody fragments binding NPN. In a typical procedure, 30,000 phage were plated and duplicate lifts with nitrocellulose screened for binding to NPN coupled to ^{125}I-labeled bovine serum albumin (BSA) (Fig. 5). Duplicate screens of 90,000 recombinant phage from the light chain library and a similar number from the heavy chain library did not identify any clones that bound the antigen. In contrast, the screen of a similar number of clones from the Fab expression library identified many phage plaques that bound NPN (Fig. 5). This observation indicates that, under conditions where many heavy chains in combination with light chains bind to antigen, heavy or light chains alone do not. Therefore, in the case of NPN, we expect that there are many heavy and light chains that only bind antigen when they are combined with specific light and heavy chains, respectively. This result supports our decision to screen large combinatorial Fab expression libraries. To assess our ability to screen large numbers of clones and obtain a more quantitative estimate of the frequency of antigen binding clones in the combinatorial library, we screened one million phage plaques and identified approximately 100 clones that bound to antigen. For six clones, a region of the plate containing the positive phage plaques and approximately 20 surrounding them was "cored," replated, and screened with duplicate lifts (Fig. 5). As expected, the expression products of approximately 1 in 20 of the phage specifically bind to antigen. Phage which were believed to be negative on the initial screen did not give positives on replating.

To determine the specificity of the antigen-antibody interaction, antigen-binding was subjected to competition with free unlabeled antigen (Fig. 6). These studies showed that individual clones could be distinguished on the basis of antigen affinity. The concentration of free haptens required for complete inhibition of binding varied between 10 to $100 \times 10^{-9} M$, suggesting that the expressed Fab fragments had binding constants in the nanomolar range.

In preparation for characterization of the protein products, a plasmid containing the heavy and light chain genes was excised with helper phage (Fig. 7). Mapping of the excised plasmid demonstrated a restriction pattern consistent with incorporation of heavy and light chain sequences. The protein products of one of the clones was

analyzed by enzyme-linked immunosorbent assay (ELISA) and immunoblotting to establish the composition of the NPN binding protein. A bacterial supernatant after IPTG (isopropyl thiogalactoside) induction was concentrated and subjected to gel filtration. Fractions in the molecular size range 40 to 60 kD were pooled, concentrated, and subjected to a further gel filtration separation. ELISA analysis of the eluted fractions (Fig. 8) indicated that NPN binding was associated with a protein of a molecular size of about 50 kD, which contained both heavy and light chains. An immunoblot of a concentrated bacterial supernatant preparation under nonreducing conditions was developed with antibody to decapeptide. This revealed a 50-kD protein band. We have found that the antigen-binding protein can be purified to homogeneity from bacterial supernate in two steps involving affinity chromatography on protein G followed by gel filtration. SDS-PAGE analysis of the protein revealed a single band at ~50 kD under nonreducing conditions and a doublet at ~25 kD under reducing conditions. Taken together, these results are consistent with NPN-binding being a function of Fab fragments in which heavy and light chains are covalently linked by a disulfide bond.

Properties of the in vivo repertoire compared to the phage combinatorial library. Previously we constructed a highly diverse V_H library in *E. coli*. We have now combined heavy and light chain libraries to clone and express assembled and functional Fab fragments of immunoglobulin. A moderately restricted library was prepared because only a limited number of primers was used for polymerase chain reaction (PCR) amplification of Fd sequences. The library is expected to contain only clones expressing κ-γ1 sequences. However, this is not an inherent limitation of the method since the addition of more primers can amplify any antibody class or subclass. Despite this restriction we were able to isolate a large number of clones producing antigen binding proteins.

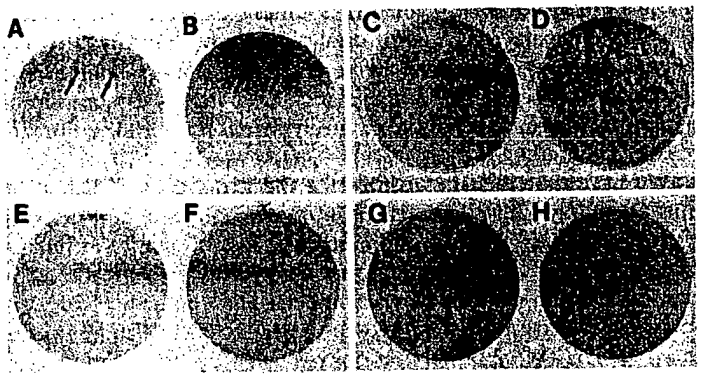


Fig. 5. Screening phage libraries for antigen-binding. Duplicate plaque lifts of Fab (filters **A** and **B**), heavy chain (filters **E** and **F**), and light chain (filters **G** and **H**) expression libraries were screened against ^{125}I-labeled BSA conjugated with NPN at a density of approximately 30,000 plaques per plate. Filters **C** and **D** illustrate the duplicate secondary screening of a cored positive from a primary filter A (arrows) as discussed in the text. Standard plaque lift methods were used in screening. Cells (XL1 blue) infected with phage were incubated on 150-mm plates for 4 hours at 37°C, protein expression was induced by overlay with nitrocellulose filters soaked in 1 m*M* IPTG, and the plates were incubated at 25°C for 8 hours. Duplicate filters were obtained during a second incubation under the same conditions. Filters were then blocked in a solution of 1 percent BSA in phosphate-buffered saline (PBS) for 1 hour before incubation (with rocking) at 25°C for 1 hour with a solution of ^{125}I-labeled BSA (at 0.1 μ*M*) conjugated to NPN (2×10^6 cpm/ml; approximately 15 NPN per BSA molecule), in 1 percent BSA in PBS. Background was reduced by preliminary centrifugation of stock ^{125}I-labeled BSA solution at 100,000*g* for 15 minutes and preliminary incubation of solutions with plaque lifts from plates containing bacteria infected with a phage having no insert. After labeling, filters were washed repeatedly with PBS containing 0.05 percent Tween 20 before the overnight development of autoradiographs.

A central issue is how our phage library compares with the in vivo antibody repertoire in terms of size, characteristics of diversity, and ease of access.

The size of the mammalian antibody repertoire is difficult to judge, but a figure of the order of 10^6 to 10^8 different antigen specificities is often quoted. With some of the reservations discussed below, a phage library of this size or larger can readily be constructed by a modification of the method described. Once an initial combinatorial library has been constructed, heavy and light chains can be shuffled to obtain libraries of exceptionally large numbers.

In principle, the diversity characteristics of the naive (unimmunized) in vivo repertoire and corresponding phage library are expected to be similar in that both involve a random combination of heavy and light chains. However, different factors act to restrict the diversity expressed by an in vivo repertoire and phage library. For example, a physiological modification such as tolerance will restrict the expression of certain antigenic specificities from the in vivo repertoire, but these specificities may still appear in the phage library. However, bias in the cloning process may introduce restrictions into the diversity of the phage library. For example, the representation of mRNA for sequences expressed by stimulated B cells can be expected to predominate over those of unstimulated cells because of higher levels of expression. In addition, the resting repertoire might overrepresent spontaneously activated B cells whose immunoglobulins have been suggested to be less specific. In any event, methods exist to selectively exclude such populations of cells. Also, the fortuitous presence of restriction sites in the variable gene similar to those used for cloning and combination will cause them to be eliminated. We can circumvent some of these difficulties by making minor changes, such as introducing amber mutations in the vector system. Different source tissues (for example, peripheral blood, bone marrow, or regional lymph nodes) and different PCR primers (for example, those to amplify different antibody classes), may result in libraries with different diversity characteristics.

Another difference between in vivo repertoire and phage library is

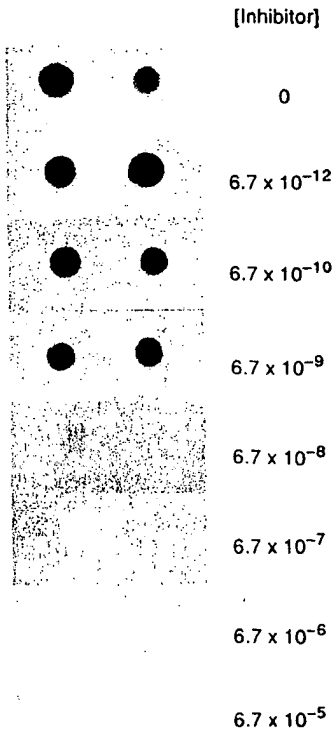


Fig. 6. Specificity of antigen binding shown by competitive inhibition. Filter lifts from positive plaques were exposed to ^{125}I-labeled BSA-NPN in the presence of increasing concentrations of the inhibitor NPN. A number of phages correlated with NPN-binding as in Fig. 5 were spotted in duplicate (about 100 particles per spot) directly onto a bacterial lawn. The plate was then overlaid with an IPTG-soaked filter and incubated for 19 hours at 25°C. The filters were then blocked in 1 percent BSA in PBS before incubation in ^{125}I-BSA–NPN as done previously with the inclusion of varying amounts of NPN in the labeling solution. Other conditions and procedures were as in Fig. 5. The results for a phage of moderate affinity are shown in duplicate in the figure. Similar results were obtained for four other phages with some differences in the effective inhibitor concentration ranges.

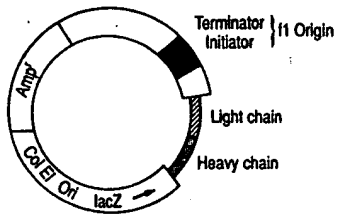


Fig. 7. A plasmid can be excised from λLc1, λHc2, and their combination because they are a modification of λzap II. M13mp8 was used as helper phage and the excised plasmid was infected into a F^+ derivative of MC1061. The excised plasmid contains the same constructs for antibody fragment expression as do the parent vectors (Fig. 1). These plasmid constructs are more conveniently analyzed for restriction pattern and protein expression of the λ phage clones identified and isolated on the basis of antigen binding. The plasmid also contains an f1 origin of replication which facilitates the preparation of single-stranded DNA for sequence analysis and in vitro mutagenesis.

that antibodies isolated from the repertoire may have benefited from affinity maturation as a result of somatic mutations after combination of heavy and light chains whereas the phage library randomly combines the matured heavy and light chains. Given a large enough phage library derived from a particular in vivo repertoire, the original matured heavy and light chains will be recombined. However, since one of the potential benefits of this technology is to obviate the need for immunization by the generation of a single highly diverse "generic" phage library, it would be useful to have methods to optimize sequences to compensate for the absence of somatic mutation and clonal selection. Three procedures are made readily available through the vector system presented. First, saturation mutagenesis may be performed on the complementarity-determining regions (CDR's) (*23*) and the resulting Fab's can be assayed for increased function. Second, a heavy or a light chain of a clone that binds antigen can be recombined with the entire light or heavy chain libraries, respectively, in a procedure identical to that used to construct the combinatorial library. Third, iterative cycles of the two above procedures can be performed to further optimize the affinity or catalytic properties of the immunoglobulin. The last two procedures are not permitted in B cell clonal selection, which suggests that the methods described here may actually increase our ability to identify optimal sequences.

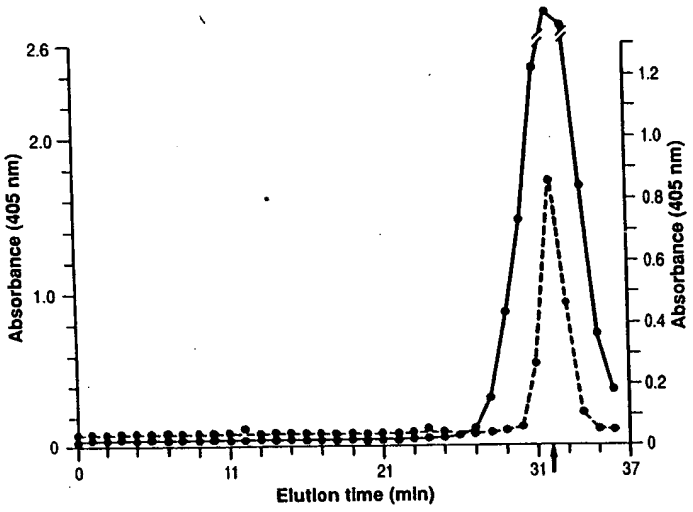


Fig. 8. Characterization of an antigen binding protein. The concentrated partially purified bacterial supernatant of an NPN binding clone was separated by gel filtration and samples from each fraction were applied to microtiter plates coated with BSA-NPN. Addition of either antibody to decapeptide (---) or antibody to κ chain (—, left-hand scale) conjugated with alkaline phosphatase was followed by color development. The arrow indicates the position of elution of a known Fab fragment. The results show that antigen binding is a property of a 50-kD protein containing both heavy and light chains. To permit protein characterization, a single plaque of a NPN-positive clone (Fig. 5) was picked, and the plasmid containing the heavy and light chain inserts (Fig. 7) was excised (*19*). Cultures (500 ml) in L broth were inoculated with 3 ml of a saturated culture of the clone and incubated for 4 hours at 37°C. Protein synthesis was induced by the addition of IPTG to a final concentration of 1 m*M*, and the cultures were incubated for 10 hours at 25°C. The supernatant from 200 ml of cells was concentrated to 2 ml and applied to a TSK-G4000 column. Samples (50 μl) from the eluted fractions were assayed by ELISA. Microtiter plates were coated with BSA-NPN at 1 μg/ml, 50-μl samples were mixed with 50 μl of PBS-Tween 20 (0.05 percent) BSA (0.1 percent) added, and the plates were incubated for 2 hours at 25°C. The plated material was then washed with PBS-Tween 20-BSA and 50 μl of appropriate concentrations of a rabbit antibody to decapeptide (*20*) or a goat antibody to mouse κ light chain (Southern Biotech) conjugated with alkaline phosphatase were added and incubated for 2 hours at 25°C. The plates were again washed, 50 μl of *p*-nitrophenyl phosphate (1 mg/ml in 0.1*M* tris, *p*H 9.5, containing 50 m*M* $MgCl_2$) was added, and the plates were incubated for 15 to 30 minutes and the absorbance was read at 405 nm.

Access is the third area where it is of interest to compare the in vivo antibody repertoire and phage library. In practical terms the phage library is much easier to access. The screening methods used have allowed one to survey the gene products of at least 50,000 clones per plate so that 10^6 to 10^7 antibodies can be readily examined in a day but the most powerful screening methods depend on selection. In the catalytic antibody system, this may be accomplished by incorporating into the antigen leaving groups necessary for replication of auxotrophic bacterial strains or toxic substituents susceptible to catalytic inactivation. Further advantages are related to the fact that the in vivo antibody repertoire can only be accessed via immunization, which is a selection on the basis of binding affinity. The phage library is not similarly restricted. For example, the only general method to identify antibodies with catalytic properties has been by preselection on the basis of affinity of the antibody to a transition state analog. Such restrictions do not apply to the in vitro library where catalysis can, in principle, be assayed directly. The ability to assay directly large numbers of antibodies for function may allow selection for catalysts in reactions where a mechanism is not well defined or synthesis of the transition state analog is difficult. Assaying for catalysis directly eliminates the bias of the screening procedure for reaction mechanisms limited to a particular synthetic analog; therefore, simultaneous exploration of multiple reaction pathways for a given chemical transformation are possible.

We have described procedures for the generation of Fab fragments that are clearly different in a number of important respects from antibodies. There is undoubtedly a loss of affinity in having monovalent Fab antigen binders, but it is possible to compensate for this by selection of suitably tight binders. For a number of applications such as diagnostics and biosensors, monovalent Fab fragments may be preferable. For applications requiring Fc effector functions, the technology already exists for extending the heavy chain gene and expressing the glycosylated whole antibody in mammalian cells.

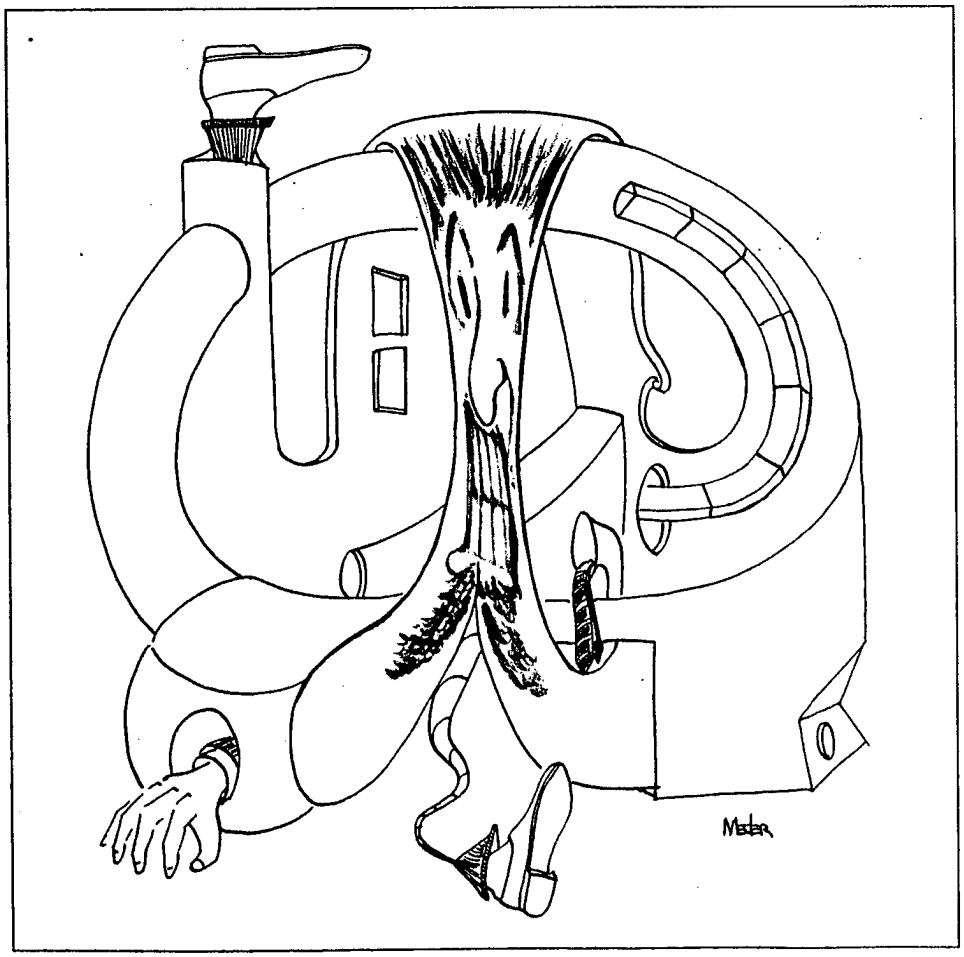
Our data show that it is now possible to construct and screen at least three orders of magnitude more clones with monospecificity than previously possible. The data also invite speculation concerning the production of antibodies without the use of live animals.

REFERENCES AND NOTES

1. R. A. Lerner, in *Proceedings of the XVIIIth Solvay Conference* on Chemistry, G. Van Binst, Ed. (Springer-Verlag, Berlin, 1983), pp. 43–49.
2. A. Tramontano, K. D. Janda, R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6736, 1986; *Science* **234**, 1566 (1986); A. Tramontano, A. A. Ammann, R. A. Lerner, *J. Am. Chem. Soc.* **110**, 2282 (1988); S. J. Pollack and P. G. Schultz, *ibid.* **111**, 1929 (1989); S. J. Pollack, P. Hsiun, P. G. Schultz, *ibid.*, p. 5961.
3. S. J. Pollack, J. W. Jacobs, P. G. Schultz, *Science* **234**, 1570 (1986); J. Jacobs, P. G. Schultz, R. Sugasawara, M. J. Powell, *J. Am. Chem. Soc.* **109**, 2174 (1987).
4. S. J. Benkovic, A. D. Napper, R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5355 (1988); K. D. Janda, R. A. Lerner, A. Tramontano, *J. Am. Chem. Soc.* **110**, 4835 (1988); K. D. Janda, S. J. Benkovic, R. A. Lerner, *Science* **244**, 437 (1989).
5. A. D. Napper, S. J. Benkovic, A. Tramontano, R. A. Lerner, *Science* **237**, 1041 (1987).

6. D. Hilvert, S. H. Carpenter, K. D. Nared, M.-T. M. Auditor, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4953 (1988); D. Y. Jackson *et al.*, *J. Am. Chem. Soc.* **110**, 4841 (1988).
7. K. D. Janda, D. Schloeder, S. J. Benkovic, R. A. Lerner, *Science* **241**, 1188 (1988). We thank K. D. Janda for providing hapten 1 and inhibitor 3.
8. A. Cochran, R. Sugasawara, P. G. Schultz, *J. Am. Chem. Soc.* **110**, 7888 (1988).
9. K. M. Shokat *et al.*, *Nature* **338**, 269 (1989).
10. B. L. Iverson and R. A. Lerner, *Science* **243**, 1184 (1989).
11. D. Hilvert *et al.*, *J. Am. Chem. Soc.*, in press.
12. R. A. Lerner and S. J. Benkovic, *Bioassays* **9**, 107 (1988); K. D. Janda, M. I. Weinhouse, D. M. Schloeder, R. A. Lerner, S. J. Benkovic, unpublished results.
13. B. L. Iverson, K. Cameron, K. Jahangiri, D. Pasternak, unpublished results.
14. L. Sastry *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5728 (1989).
15. E. S. Ward, D. Gussow, A. D. Griffiths, P. T. Jones, G. Winter, *Nature* **341**, 544 (1989); R. Orlandi, D. H. Gussow, P. T. Jones, G. Winter, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 3833 (1989).
16. D. R. Davies, S. Sheriff, E. A. Padlan, *J. Biol. Chem.* **263**, 10541 (1988); J. N. Herron, X. He, M. L. Mason, E. W. Voss, Jr., A. B. Edmundson, *Proteins Struct. Funct. Genet.* **5**, 271 (1989); M. Whitlow, A. J. Howard, J. F. Wood, E. W. Voss, Jr., K. D. Hardman, unpublished results. In the structure of Whitlow *et al.*, the light chain in the fluorescein isothiocyanate conjugated Fab 4-4-20 crystal structure has major contributions to the binding of the hapten fluorescein. All of the hydrogen bonds to fluorescein are from the light chain. Three come from the light chain CDR 1 (residues His^{27D}, Tyr^{32}, and Arg^{34}, in Kabat's numbering system) and one from the light chain CDR 3 (residue Ser^{91}). Fluorescein is bound in a deep pocket in the Fab 4-4-40 structure, so that only 10 percent (54 Å^2 of 521 Å^2) of its solvent accessible surface area is exposed to solvent; 38 percent of this pocket is lined with light chain residues, based on calculations of solvent-accessible surface area. Aromatic residues His^{27D}, Tyr^{32}, and Trp^{96} from the light chain and Trp^{33}, Tyr^{53}, Tyr^{100D}, and Tyr^{100E} from the heavy chain form this deep pocket for the fluorescein.
17. S. C. Wallick, E. A. Kabat, S. L. Morrison, *Immunology* **142**, 1235 (1989).
18. M. Better, C. P. Chang, R. R. Robinson, A. H. Horwitz, *Science* **240**, 1041 (1988); A. Skerra and A. Plückthun, *ibid.*, p. 1038.
19. J. M. Short, J. M. Fernandez, J. A. Sorge, W. D. Huse, *Nucleic Acids Res.* **16**, 7583 (1988).
20. J. Field *et al.*, *Mol. Cell. Biol.* **8**, 2159 (1988).
21. R. Orlandi, D. H. Gussow, P. T. Jones, G. Winter, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 3833 (1989).
22. The Fv (variable region) fragments may be constructed with a 3′ primer that is complementary to the mRNA in the J (joining) region (amino acid 128) and a set of 5′ primers that are complementary to the first strand cDNA in the conserved amino-terminal region of the processed protein.
23. Amino acid and nucleotide sequences of immunoglobulins used in this paper were taken from E. A. Kabat, T. T. Wu, M. Reid-Miller, H. M. Perry, K. S. Gottesman, *Sequences of Proteins of Immunological Interest* (US Public Health Service, National Institutes of Health, Bethesda, MD, 1987).
24. We thank D. Schloeder, D. A. McLeod, R. Samodal, B. Hay for technical assistance; E. A. Kabat, N. Klinman, K. D. Janda, and B. L. Iverson for advice and comments. Supported by an NIH postdoctoral fellowship (S.A.I), the SERC Protein Engineering Club (A.K.), and a Jenner Fellowship of the Lister Institute of Preventive Medicine (D.R.B.).

26 October 1989; accepted 14 November 1989



" Convinced of the correctness of his 17-dimensional model of the universe, theorist Martin Nowsky was not above testing it directly."

C14

Antibodies Without Immunization

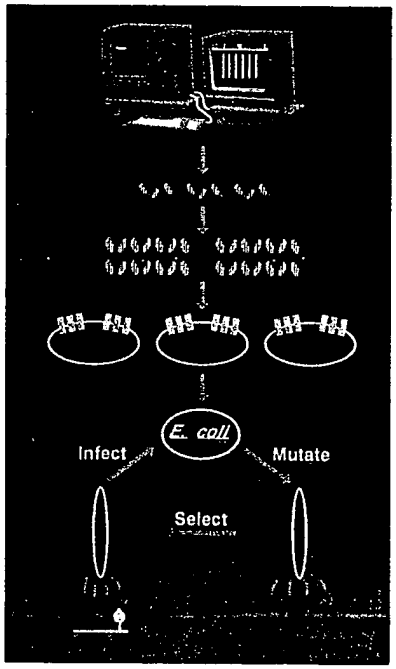
Richard A. Lerner, Angray S. Kang, Jim D. Bain, Dennis R. Burton, Carlos F. Barbas, III

Immunization of animals to produce antibodies for laboratory and medical use is among the time-honored processes in science. However, recent experimental data have indicated that this process might be replaced with a synthetic one. Examination of how naturally occurring antibodies are formed reveals some critical features and mechanisms that should be considered during the design of a replacement process: (i) the immune system is combinatorial, (ii) the processes of replication and recognition are linked, (iii) the antibody gene pool is largely naïve, and (iv) the system is capable of somatic refinement.

Copying and even improving this natural complexity requires the replication of the above features in a set of linked transactions. The ultimate goal is to provide a library of antibodies with nearly endless binding specificities and without dependency on animals or their cells.

The natural antibody system is fundamentally combinatorial. Diversity is generated by combinations of genes encoding two presumably randomly associated protein chains (heavy and light), each of which contains three hypervariable or complementarily determining regions (CDR's). The vast diversity of the antibody repertoire is achieved by the combinatorial association of these six CDR's in three-dimensional space. In a synthetic system the main hurdle to replacing the process of immunization is the need to mimic in vitro this large combinatorial process. This task is the opposite to that presented by cloning because diversity rather than homogeneity must be achieved. A further complication in antibody synthesis is that the CDR's are present on two different protein chains.

Earlier work on the assembly of immunoglobulins in *Escherichia coli* (*1–3*) and on amplification of the individual genes by the polymerase chain reaction (PCR) (*4–6*) provided a basis for the in vitro generation of antibody diversity. However, progress in the field was hampered by the assumption on the part of investigators that the probability of obtaining useful antibodies by random chain association was too low. Nevertheless, it has now been shown that the combinatorial problem could be solved by random association and expression of large numbers of genes encoding heavy and light chains in phage systems (*7*). This method is efficient for immunized animals because, as exposure to antigen continues, the induced repertoire becomes more homogeneous and results in an increased concentration of chains that can be paired in a way appropriate for antigen binding (*8*). Further, it may not be necessary to recapitulate the chain pairing found in the immunized animal. In fact, diversity can often be increased by a process of chain shuffling (*9*). A solution to the combinatorial problem cleared the main obstacle to replacing immunization in that an antibody library with a potential diversity exceeding that of intact animals could be contained in a volume of only a few microliters. It then remained necessary to obtain the starting gene pool, access large libraries, and improve their binding specificities.



Synthetic generation of antibodies. Sequences for antibody chains are constructed from synthetic nucleic acids randomized in the antigen binding region. These sequences are cloned into bacteriophages, which are then subject to mutation and selection.

Biological systems are especially powerful because recognition and replication are linked. Linkage can be at the level of whole organisms, individual cells, or viral entities. In the immune system the linkage exists because cells that recognize antigens at their surfaces are stimulated to divide. Screening of combinatorial antibody libraries is aimed at approximating this central feature of biological systems.

Systems in which diverse peptide sequences are displayed on the surface of filamentous bacteriophages (*10, 11*) have proved attractive for linking combinatorial antibodies with the genes that encode them in a package capable of recognizing antigen and replicating. This has been done in two ways (*12–14*). In each case both antibody chains (heavy and light) are linked to gene fragments that encode signals that direct them to the periplasmic space of *E. coli*. The genes for the antibody heavy chains are fused with either the phage gene III or gene VIII, which encode low and high copy phage surface proteins, respectively. This design allows heavy chains linked to phage proteins to encounter free light chains in the periplasmic space during the assembly of phage particles. Since the volume of an *E. coli* bacterium is somewhat less than 10^{-15} liter, a few hundred molecules of the antibody chains in the periplasmic space exceed the association constant of the heterodimer, thereby leading to appropriate assembly of the two-chain antibody structure on the phage surface. The difference between the methods is that in one the antibody genes are incorporated into the phage genome whereas in the other method they are in a smaller gene (phagemid), which is copied during viral replication. We use the phagemid (*13, 14*) because it readily permits gene cloning and its small size allows the generation of very large libraries. Furthermore, the phagemid–gene III constructions have yielded monovalent phage, thereby facilitating the selection of the highest affinity antibody clones.

Phage with specific antigen-binding properties can be selected and enriched by repeated attachment to, and elution from, a source of immobilized antigen. Although a concentration of 10^{11} phage per 50 microliters is possible, the transforming efficiency of the DNA limits the process so that in practice only 10^8 different antibodies can be readily studied. Nevertheless, with these methods of screening combinatorial libraries, many antibodies to haptens and proteins from immunized animals and man have already been prepared.

With combinatorial libraries available, attention turned to circumventing the process of immunization. A primary consideration was the source of the antibody genes that make up the library (*15*). If educated

R. A. Lerner, A. S. Kang, J. D. Bain, and C. F. Barbas, III, are in the Departments of Chemistry and Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037. D. R. Burton is in the Department of Immunology, The Scripps Research Institute, La Jolla, CA 92037.

libraries are used, the investigator is engaged in the formidable process of protein remodeling. Imagine the attendant difficulty of selecting a chain combination that must bind a hydrophobic hapten deeply within a binding pocket from a pool of antibody chains largely evoked to maximize the area of interacting protein surfaces to bind a protein antigen.

Naïve combinatorial libraries require a source of genes. Immunoglobulin genes from animals do not constitute a naïve gene pool because the animals have received many antigenic challenges. Furthermore, an animal repertoire may be evolutionarily restricted. For example, evidence is emerging that in humans the use of antibody chains is far from random. A further complication is the likely need for PCR to amplify sequences from natural sources because diversity decreases on subsequent rounds of gene amplification, a situation that may be favorable in immunized but not unimmunized animals.

Synthesizing combinatorial libraries then became the method of choice. Here some or all of the CDR's are encoded by synthetic nucleic acids randomized at most positions (*16*). Such libraries have a potential diversity greater than that of animals. We have selected high-affinity antibodies to fluorescein from combinatorial libraries in which the CDR3 regions of both the light and heavy chains are encoded by synthetic gene segments (*16*). Analysis of codon usage showed that selection for binding was attributable to the protein and was not due to some bias either in nucleic acid synthesis or in clonal expansion during phage selection (*16*). The selected antibodies showed consensus sequences related to both the general structural requirements of the antibody molecule and features required for specific binding to fluorescein. For example, only antibodies that had aspartic acid in position 101 of the heavy chain CDR3 were selected (*16*). In naturally occurring antibodies, the carboxylate of Asp^{101} often has an important structural role in that it can form a salt bridge with the guanidinium moiety of a framework arginine residue, thereby constraining the conformation of the CDR's (*17*). Thus, the in vitro selection system is mimicking a feature that results from natural selection and showed that antibody chains with synthetic CDR's can form the basis of large combinatorial libraries. Although selection for general structural features of the antibody molecule occurred when mixtures of synthetic and natural CDR's were combined, even this requirement may be relaxed when all the CDR's are synthetic. Thus, the system may have enough degrees of freedom to generate useful CDR combinations in a way not possible unless all the combinatorial parameters can vary in concert.

In some instances it will be necessary to mimic the natural somatic refinement process to increase the affinity of antibodies selected in the primary screening. We have been able to select for mutated antibodies of higher affinity by iterative procedures in which mutations are introduced by error-prone PCR (*18*). Although this staggered process of mutation and selection differs from the natural integrated selection scheme for cells with the highest binding affinity, the number of independent binding events that are available for study may be higher than those of natural systems. Whether sorting through large numbers of binding events can approximate or even exceed affinities obtained by more tightly integrated mutation and selection remains to be seen.

Apparently there is no longer any major technical obstacle to eliminating the process of immunization or even the use of animals for the preparation of antibodies. Whether or not the technology is general and applicable to all classes of molecules, we do know that principles learned for haptens have been generalizable to protein antigens. We anticipate that soon master combinatorial antibody libraries will be available for distribution and that such libraries will depend on the use of synthetic CDR's in which all positions are randomized except those known to be conserved in natural antibodies. Also, the length of the CDR's will be varied and will include some that are longer than those of natural antibodies. Only for human antibodies intended for therapy need considerations of potential immunogenicity temper any search for novel structures.

Combinatorial antibody libraries should be useful for isolating large enough numbers of antibodies to search for consensus sequence motifs implicated in binding to a particular class of antigens. This approach appears promising for generating antibodies that neutralize HIV. Here, highly specialized libraries may be constructed where motifs known to affect virus neutralization are incorporated into the synthetic CDR's. The master library should grow by acquiring these specialized libraries. Hence, the scientific community should adopt a technology that is sufficiently uniform for new constructs to be easily incorporated into the master library.

The use of master libraries could be compared to the process of immunization. The initial selection is like the primary response in which the affinity is expected to be variable and highly dependent on the chemical nature of the antigen. If the affinity constant is not suitable, mutagenesis and reselection, a process related to somatic refinement during sustained immunization, could be instituted. The iterative selection procedures used for screening and even mutating combinatorial libraries lend themselves to automation, and we can expect to see machines that carry out many of the steps.

Combinatorial antibody libraries were generated initially to match the chemical component of antibody catalysis with a protein component that was more understandable in chemical terms. Synthetic combinatorial antibody libraries have a diversity that is not subject to in vitro editing processes such as tolerance. This last feature allows selection for rare antibody species like those that may be required for acid-base catalysis or those that operate on structures of low immunogenicity, such as nucleic acids. In addition, unlike induction in animals that selects for binding events under aqueous conditions and neutral pH, antibodies can be selected from synthetic libraries that bind under various reaction conditions, including in the presence of organic solvents. Even antibodies that make covalent interactions can be selected by elution with nucleophiles. When known protein binding motifs are used, antibody chains that incorporate cofactors and metals can be included in the combinatorial library and thus extend the range of antibody catalysis to chemical reactions beyond the capability of amino acid side chains. Finally, master libraries can be thought of as a uniform and reproducible source of antibodies, a feature that can increase the precision of some immunological techniques and at the same time be of service to chemistry, biology, and medicine.

REFERENCES AND NOTES

1. M. A. Boss, J. H. Kenten, C. R. Wood, J. S. Emtage, *Nucleic Acids Res.* 12, 3791 (1984).
2. S. Cabilly *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 81, 3273 (1984).
3. M. Better, C. P. Chang, R. R. Robinson, A. H. Horwitz, *Science* 240, 1041 (1988).
4. J. W. Larrick, Y. L. Chiang, R. Sheng-Dong, G. Senyk, P. Casali, *Prog. Biotechnol.* 5, 231 (1988).
5. R. Orlandi, D. H. Gussow, P. T. Jones, G. Winter, *Proc. Natl. Acad. Sci. U.S.A.* 86, 3833 (1989).
6. L. Sastry *et al.*, *ibid.*, p. 5728.
7. W. D. Huse *et al.*, *Science* 246, 1275 (1989).
8. D. R. Burton, *Trends Biotechnol.* 9, 169 (1991).
9. A. S. Kang, T. M. Jones, D. R. Burton, *Proc. Natl. Acad. Sci. U.S.A.* 88, 11120 (1991).
10. J. K. Scott and G. P. Smith, *Science* 249, 386 (1990).
11. S. E. Cwirla, E. A. Peters, R. W. Barrett, W. J. Dower, *Proc. Natl. Acad. Sci. U.S.A.* 87, 6378 (1990).
12. J. McCafferty, A. D. Griffiths, G. Winter, D. J. Chiswell, *Nature* 348, 552 (1990).
13. A. S. Kang, C. F. Barbas, III, K. D. Janda, S. J. Benkovic, R. A. Lerner, *Proc. Natl. Acad. Sci U.S.A.* 88, 4363 (1991).
14. C. F. Barbas, III, A. S. Kang, R. A. Lerner, S. J. Benkovic, *ibid.*, p. 7978.
15. G. Winter and C. Milstein, *Nature* 349, 293 (1991).
16. C. F. Barbas, III, J. D. Bain, D. M. Hoekstra, R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.* 89, 4457 (1992).
17. C. Chothia and A. M. Lesk, *J. Mol. Biol.* 196, 901 (1987).
18. H. Gram *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 89, 3576 (1992).

J. Mol. Biol. (1991) 222, 581–597

C17

By-passing Immunization

Human Antibodies from V-gene Libraries Displayed on Phage

James D. Marks[1], Hennie R. Hoogenboom[1], Timothy P. Bonnert[1]
John McCafferty[3], Andrew D. Griffiths[1] and Greg Winter[1,2]†

[1]*MRC Centre for Protein Engineering and*
[2]*MRC Laboratory of Molecular Biology*
Hills Road, Cambridge CB2 2QH, U.K.

[3]*Cambridge Antibody Technology Ltd*
Daly Research Laboratories
Babraham Hall, Cambridge CB2 4AT, U.K.

(Received 6 September 1991; accepted 27 September 1991)

We have mimicked features of immune selection to make human antibodies in bacteria. Diverse libraries of immunoglobulin heavy (V_H) and light (V_κ and V_λ) chain variable (V) genes were prepared from peripheral blood lymphocytes (PBLs) of unimmunized donors by polymerase chain reaction (PCR) amplification. Genes encoding single chain Fv fragments were made by randomly combining heavy and light chain V-genes using PCR, and the combinatorial library ($>10^7$ members) cloned for display on the surface of a phage. Rare phage with "antigen-binding" activities were selected by four rounds of growth and panning with "antigen" (turkey egg-white lysozyme (TEL) or bovine serum albumin) or "hapten" (2-phenyloxazol-5-one (phOx)), and the encoding heavy and light chain genes were sequenced. The V-genes were human with some nearly identical to known germ-line V-genes, while others were more heavily mutated. Soluble antibody fragments were prepared and shown to bind specifically to antigen or hapten and with good affinities, K_a (TEL) $= 10^7$ M^{-1}; K_a (phOx) $= 2 \times 10^6$ M^{-1}. Isolation of higher-affinity fragments may require the use of larger primary libraries or the construction of secondary libraries from the binders. Nevertheless, our results suggest that a single large phage display library can be used to isolate human antibodies against any antigen, by-passing both hybridoma technology and immunization.

Keywords: filamentous phage; human antibodies; combinatorial libraries

1. Introduction

Over the last century animal antiserum, and more recently rodent monoclonal antibodies, have been used clinically to neutralize toxins, and to treat bacterial and viral infections. In the future the specific recognition of human cell-surface markers by antibody fragments should enable functional manipulations of subsets of immuno-competent cells in the fields of, for example, autoimmunity, transplantation, and the inhibition of cell adhesion and of cytokine-stimulated cell proliferation. However, the use of animal antibody can lead to an antiglobulin response and hypersensitivity reactions. Ideally human monoclonal antibodies would be used, but it is difficult to make them. Not only are peripheral blood lymphocytes (PBLs‡) a poor source of the blast cells that are actively involved in the immune response, but it is difficult to immortalize them. The use of mouse myeloma lines as fusion partners for human B-cells leads to a preferential loss of human chromosomes and instability of the hybrids, and Epstein Barr virus infection of B-cells also tends to

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: PBL, peripheral blood lymphocyte; Ig, immunoglobulin; PCR, polymerase chain reaction; g3p, gene *3* protein; ELISA, enzyme-linked immunosorbent assay; BSA, bovine serum albumin; TEL, turkey egg-white lysozyme; t.u., transducing unit(s); p.f.u., plaque-forming units(s); IPTG, isopropyl β-D-thiogalactopyranoside.

produce unstable (IgM) lines with poor antigen affinity (for a review and references, see Winter & Milstein (1991)).

However, there are other ways of tapping the antibody repertoire of immunized humans or animals. Instead of immortalizing B-cells for production of monoclonal antibodies, the antibody heavy and light chain V-genes are immortalized by gene technology, and antibodies or fragments expressed in mammalian cells, yeast or bacteria. For example, recombinant antibodies were rescued from hybridomas by PCR amplification of the V-genes with "universal" primers, and cloning the genes into vectors for expression of complete antibodies (Orlandi *et al.*, 1989). In principle this technique could be extended to the construction of antibodies from the V-genes of single B-cells, thereby bypassing hybridoma technology (Orlandi *et al.*, 1989; Larrick *et al.*, 1989). Alternatively, libraries of V-genes have been used to express soluble antibody fragments, which are then screened for antigen-binding activities (Ward *et al.*, 1989; Huse *et al.*, 1989; Caton & Koprowski, 1990; Mullinax *et al.*, 1990; Persson *et al.*, 1991). For example, from a donor immunized with tetanus toxoid, V-genes from the mRNA of 10^8 human PBLs were combined at random in bacteriophage lambda, so scrambling the original heavy and light chain pairings. When the combinatorial library (10^7 members) was expressed in bacteria and 12,000 plaques were screened on nitrocellulose filters for binding to toxoid, 10 binders were found (Mullinax *et al.*, 1990). Thus, human antibodies can be made by filter screening of combinatorial libraries from immunized donors.

By contrast we have avoided the screening of large numbers of individual clones on filters by mimicking features of immune selection (Milstein, 1990; McCafferty *et al.*, 1990; Winter & Milstein, 1991). In the immune system, diverse combinatorial libraries of antibodies are displayed on the surface of B-cells, and specific recognition with antigen triggers cell proliferation and differentiation into antibody-secreting or memory pathways. We have displayed (Smith, 1985; Parmley & Smith, 1988) antibody fragments on the surface of filamentous bacteriophage by fusion to a minor coat protein at the tip of the phage, the gene *3* protein (g3p) (McCafferty *et al.*, 1990). Phage encoding antibody fragments with binding activities were selected from those encoding non-binders by affinity chromatography. By rounds of growth and selection, rare binders were selected, with an enrichment of one in 10^3 after one round of panning, and one in 10^6 after two rounds (McCafferty *et al.*, 1990). Antibody fragments can be displayed as fusions with g3p as single polypeptide chains in which the heavy and light chain variable domains are linked by a polypeptide spacer (single chain Fv or scFv: McCafferty *et al.*, 1990), or as non-covalently associated heavy and light chains (Fab fragments) (Hoogenboom *et al.*, 1991). Fab fragments have also been displayed as fusions with the major coat protein (gene *8*: Kang *et al.*, 1991). Recently we used phage to display a small random combinatorial library (2×10^5 members) of scFv antibody fragments from the spleen mRNA of immunized mice (Clackson *et al.*, 1991). The mRNA is presumably derived mainly from plasma cells (R. Hawkins & G. Winter, unpublished results), as the level of Ig mRNA in these cells is up to 1000-fold greater than in resting B-cells (Schibler *et al.*, 1978). After only a single round of affinity selection, we isolated numerous different antibodies with affinities in the range of 10^5 M^{-1} to 10^8 M^{-1}.

However, it is rarely possible to immunize humans to order, and the possibility of making human antibodies without prior immunization is particularly appealing. We have therefore applied the phage display technology to making human antibodies from V-gene repertoires from unimmunized donors. We made a large scFv library from the PBLs, and with greater than 10^7 members it was similar in size to the B-cell repertoire of a mouse at any one moment. The library was also made as diverse as possible by using both V_κ and V_λ light chains, as well as V_Hs derived from IgM and IgG mRNA. Diversity was further maximized by using PCR primers based on each of the human heavy and light chain gene families (Marks *et al.*, 1991). Finally, the library was subjected to multiple rounds of affinity selection to ensure that even a single clone in the original library could be isolated.

2. Materials and Methods

(a) *Primer design*

We optimized the design of the PCR primers for the rearranged V-genes to maximize the diversity of the PCR products. The primers were located at the 5′ and 3′ ends (back and forward primers, respectively) of the mature V-regions (Orlandi *et al.*, 1989; Marks *et al.*, 1991; Songsivilai *et al.*, 1990), but did not incorporate internal restriction sites that mismatch the template and bias amplification. The back primers were designed to match each of the families of human V-genes, and forward primers to match each of the human germ-line J-segments (Table 1). Furthermore, sets of PCR primers were designed to optimize the linking of V_H and V_κ or V_λ genes at random, and append restriction sites to the linked genes (Table 1 and Fig. 1).

(b) *Assay of donor serum for presence of IgM antibodies to phOx-BSA and TEL*

Serum from the 2 donors was assayed for the presence of IgM antibodies to phOx-BSA and TEL using an ELISA-based assay kit for detection of human IgM antibodies in serum (Platest, Menarini Diagnostics). Microtiter plates were coated overnight with either 10 μg phOx-BSA/ml or 10 μg TEL/ml. Plates were washed 3 times with PBS (phosphate-buffered saline: 25 mM-NaH_2PO_4, 125 mM-NaCl, pH 7·0) and blocked for 2 h with 2% MPBS (2% (w/v) skimmed milk powder (Marvel) in PBS) at 37°C. Donor serum was diluted 1/40 in PBS and 50 μl was added to the microtiter wells and incubated for 30 min at room temperature. The plates were washed 3 times with PBS and 50 μl horseradish peroxidase-conjugated anti-human IgM antibody was

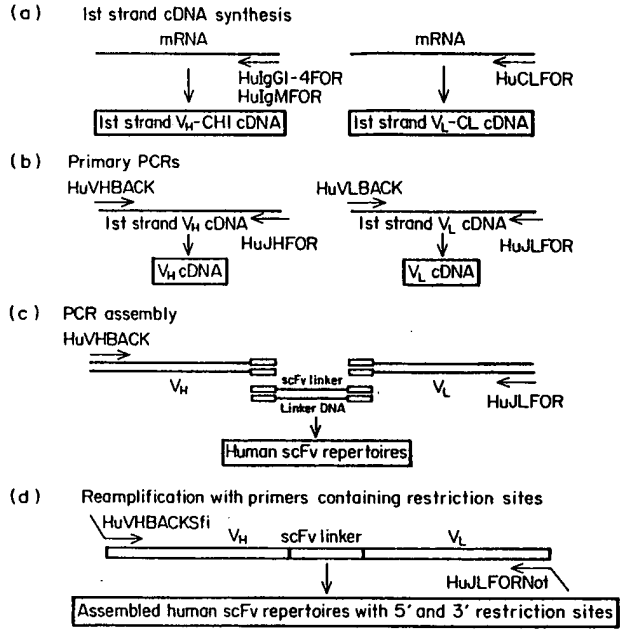


Figure 1. Making scFv gene repertoires. (a) mRNA is primed with constant region-specific oligonucleotides and 1st strand cDNA synthesized. (b) Portions of 1st strand cDNA are PCR amplified with a mixture of V-gene and J-segment primers. (c) The rearranged V_H and V_L PCR products are combined in a 2nd PCR amplification containing linker DNA that overlaps the C terminus of the V_H and the N terminus of the V_L genes. This reaction mixture is subjected to temperature cycling followed by amplification. (d) Finally, the resulting scFv gene repertoires are reamplified with primers containing appended restriction sites.

added to each well and incubated for 30 min. Plates were washed 3 times with PBS, developed as in the kit protocol and the plate read at 450 nm.

(c) *cDNA synthesis, PCR amplication and assembly of scFv genes*

Blood (500 ml) containing approximately 10^8 B-lymphocytes, was obtained from 2 healthy volunteers. The white cells were separated in Ficoll and RNA was prepared using a modified method described by Cathala *et al.* (1983). Heavy chain repertoires were prepared from both IgG and IgM cDNA in order to tap both mature and naive lymphocytes (Roit *et al.*, 1985), and light chain repertoires were prepared from both V_κ and V_λ genes. Thus, 4 first strand cDNA syntheses were made as described (Marks *et al.*, 1991) from RNA corresponding to $2{\cdot}5 \times 10^7$ B-cells, using either an IgG or an IgM constant region primer for the heavy chains, or a κ or λ constant region primer for light chains (Table 1 and Fig. 1(a)). All of the cDNA was used to generate 4 separate repertoires of scFv genes ($V_{H\mu}$-V_κ, $V_{H\mu}$-V_λ, $V_{H\gamma}$-V_κ $V_{H\gamma}$-V_λ) as described below (Figs 1 and 2).

V_H, V_κ and V_λ-genes were amplified separately using an equimolar mixture of the appropriate family-based back and forward primers (Table 1, Figs 1(b) and 2). Reaction mixtures (50 μl) were prepared containing 5 μl of the supernatant from the cDNA synthesis, 20 pmol back primers, 20 pmol forward primers, 250 μM-dNTPs 10 mM-KCl, 10 mM-$(NH_4)_2SO_4$, 20 mM-Tris·HCl (pH 8·8), 2·0 mM-$MgCl_2$, 100 μg BSA/ml and 1 μl (1 unit) Vent DNA polymerase (New England Biolabs). The reaction mixture was overlaid with mineral (paraffin) oil and subjected to 30 cycles of amplification using a Techne thermal cycler. The cycle was 94°C for 1 min (denaturation), 57°C for 1 min (annealing) and 72°C for 1 min (extension). The products were purified on a 2% (w/v) agarose gel, isolated from the gel by Geneclean (Bio-101) and resuspended in 25 μl of water.

To make the scFv linker DNA, 52 separate 50 μl PCR reactions were performed using each of the 4 reverse JH primers in combination with each of the 13 reverse V_κ and V_λ oligonucleotides (Table 1). The template was approximately 1 ng of pSW2scFvD1·3 (McCafferty *et al.*, 1990) containing the short peptide $(Gly_4Ser)_3$ (Huston *et al.*, 1988). The PCR reaction reagents were as described above and the cycle was 94°C for 1 min, 45°C for 1 min and 72°C for 1 min. The linkers were purified on a 2% agarose gel, eluted from the gel on a Spin-X column (Costar) and precipitated with ethanol.

For PCR assembly of the scFv repertoires (Fig. 1(c)), approximately 1 μg of a primary heavy chain amplification ($V_{H\mu}$ or $V_{H\gamma}$) and 1 μg of a primary light chain amplification (V_κ or V_λ) were combined with approximately 250 ng of the appropriate linker DNA (an equimolar mixture of each of the 6 JH-V_κ or 7 JH-V_λ linkers) in a 50 μl PCR reaction mixture and cycled 7 times (94°C for 2 min and 72°C for 2·5 min) to join the fragments. The reaction mixture was then amplified for 25 cycles (94°C for 1 min and 72°C for 3 min) after the addition of 20 pmol of the outer PCR primers (Fig. 1(c)). Finally, the assembled products were gel-purified and reamplified for 25 cycles (94°C for 1 min, 55°C for 1 min, 72°C for 2·5 min) with the flanking oligonucleotides containing the appended restriction sites (Fig. 1(d)). PCR buffers and dNTPs were as described previously. The resulting scFv repertoires ($V_{H\mu}$-V_κ, $V_{H\mu}$-V_λ, $V_{H\gamma}$-V_κ, $V_{H\gamma}$-V_λ were purified on a 1·5% agarose gel, electroeluted and precipitated with ethanol (Sambrook *et al.*, 1990). For subsequent cloning, the $V_{H\mu}$-V_κ and $V_{H\mu}$-V_λ repertoires were combined (IgM repertoire) as were the $V_{H\gamma}$-V_κ and $V_{H\gamma}$-V_λ repertoires (IgG repertoire).

(d) *Cloning of the scFv gene repertoires*

Purified DNA of the scFv gene repertoires (1 to 4 μg) was digested with *Not*I and either *Sfi*I or *Nco*I restriction enzymes. (The 2 different restriction enzymes were tried in an attempt to increase the cloning efficiency.) After digestion, the fragments were extracted with phenol/chloroform, and ligated into pHEN1 (Hoogenboom *et al.*, 1991) vector that had been digested with either *Sfi*I and *Not*I or *Nco*I and *Not*I and electroeluted from a 0·8% agarose gel (Sambrook *et al.*, 1990). Each scFv gene repertoire was combined in a ligation mixture which included 6 μg of digested vector, in a 100 μl ligation mix with 2000 units of phage T4 DNA ligase (New England Biolabs) overnight at room temperature. The ligation mix was purified by extraction with phenol and precipitation with ethanol. The ligated DNA was resuspended in 10 μl of water, and 2·5 μl samples were electroporated (Dower *et al.*, 1988) into 50 μl *Escherichia coli* TG1 (Gibson, 1984). Cells were grown in 1 ml of SOC (Sambrook *et al.*, 1990) for 1 h and then plated on TYE (Miller, 1972) medium with 100 μg ampicillin/ml and 1% (w/v) glucose (TYE-AMP-GLU), in 243 mm × 243 mm dishes (Nunc). Colonies were scraped off the plates into 10 ml of 2×TY broth (Miller, 1972) containing 100 μg ampicillin/ml. 1%

Table 1

Oligonucleotide primers used for PCR of human immunoglobulin genes

A. *1st strand cDNA synthesis*

Human heavy chain constant region primers

HuIgG1-4CH1FOR	5'-GTC CAC CTT GGT GTT GCT GGG CTT-3'
HuIgMFOR	5'-TGG AAG AGG CAC GTT CTT TTC TTT-3'

Human κ constant region primer

HuGκFOR	5'-AGA CTC TCC CCT GTT GAA GCT CTT-3'

Human λ constant region primer

HuCλFOR	5'-TGA AGA TTC TGT AGG GGC CAC TGT CTT-3'

B. *Primary PCRs*

Human V_H back primers

HuVH1aBACK	5'-CAG GTG CAG CTG GTG CAG TCT GG-3'
HuVH2aBACK	5'-CAG GTC AAC TTA AGG GAG TCT GG-3'
HuVH3aBACK	5'-GAG GTG CAG CTG GTG GAG TCT GG-3'
HuVH4aBACK	5'-CAG GTG CAG CTG CAG GAG TCG GG-3'
HuVH5aBACK	5'-GAG GTG CAG CTG TTG CAG TCT GC-3'
HuVH6aBACK	5'-CAG GTA CAG CTG CAG CAG TCA GG-3'

Human J_H forward primers

HuJH1-2FOR	5'-TGA GGA GAC GGT GAC CAG GGT GCC-3'
HuJH3FOR	5'-TGA AGA GAC GGT GAC CAT TGT CCC-3'
HuJH4-5FOR	5'-TGA GGA GAC GGT GAC CAG GGT TCC-3'
HuJH6FOR	5'-TGA GGA GAC GGT GAC CGT GGT CCC-3'

Human V_κ back primers

HuVκ1aBACK	5'-GAC ATC CAG ATG ACC CAG TCT CC-3'
HuVκ2aBACK	5'-GAT GTT GTG ATG ACT CAG TCT CC-3'
HuVκ3aBACK	5'-GAA ATT GTG TTG ACG CAG TCT CC-3'
HuVκ4aBACK	5'-GAC ATC GTG ATG ACC CAG TCT CC-3'
HuVκ5aBACK	5'-GAA ACG ACA CTC ACG CAG TCT CC-3'
HuVκ6aBACK	5'-GAA ATT GTG CTG ACT CAG TCT CC-3'

Human J_κ forward primers

HuJκ1FOR	5'-ACG TTT GAT TTC CAC CTT GGT CCC-3'
HuJκ2FOR	5'-ACG TTT GAT CTC CAG CTT GGT CCC-3'
HuJκ3FOR	5'-ACG TTT GAT ATC CAC TTT GGT CCC-3'
HuJκ4FOR	5'-ACG TTT GAT CTC CAC CTT GGT CCC-3'
HuJκ5FOR	5'-ACG TTT AAT CTC CAG TCG TGT CCC-3'

Human λ back primers

Huλ1BACK	5'-CAG TCT GTG TTG ACG CAG CCG CC-3'
Huλ2BACK	5'-CAG TCT GCC CTG ACT CAG CCT GC-3'
Huλ3aBACK	5'-TCC TAT GTG CTG ACT CAG CCA CC-3'
Huλ3bBACK	5'-TCT TCT GAG CTG ACT CAG GAC CC-3'
Huλ4BACK	5'-CAC GTT ATA CTG ACT CAA CCG CC-3'
Huλ5BACK	5'-CAG GCT GTG CTC ACT CAG CCG TC-3'
Huλ6BACK	5'-AAT TTT ATG CTG ACT CAG CCC CA-3'

Human λ forward primers

HuJλ1FOR	5'-ACC TAG GAC GGT GAC CTT GGT CCC-3'
HuJλ2-3FOR	5'-ACC TAG GAC GGT CAG CTT GGT CCC-3'
HuJλ4-5FOR	5'-ACC TAA AAC GGT GAG CTG GGT CCC-3'

C. *PCR assembly*

Reverse J_H for scFv linker

RHuJH1-2	5'-GCA CCC TGG TCA CCG TCT CCT CAG GTG G-3'
RHuJH3	5'-GGA CAA TGG TCA CCG TCT CTT CAG GTG G-3'
RHuJH4-5	5'-GAA CCC TGG TCA CCG TCT CCT CAG GTG G-3'
RHuJH6	5'-GGA CCA CGG TCA CCG TCT CCT CAG GTG C-3'

Reverse V_κ for scFv linker

RHuVκ1aBACKFv	5'-GGA GAC TGG GTC ATC TGG ATG TCC GAT CCG CC-3'
RHuVκ2aBACKFv	5'-GGA GAC TGA GTC ATC ACA ACA TCC GAT CCG CC-3'
RHuVκ3aBACKFv	5'-GGA GAC TGC GTC AAC ACA ATT TCC GAT CCG CC-3'
RHuVκ4aBACKFv	5'-GGA GAC TGG GTC ATC ACG ATG TCC GAT CCG CC-3'
RHuVκ5aBACKFv	5'-GGA GAC TGC GTG AGT GTC GTT TCC GAT CCG CC-3'
RHuVκ6aBACKFv	5'-GGA GAC TGA GTC AGC ACA ATT TCC GAT CCG CC-3'

Table 1 *(continued)*

Reverse V_λ for scFv linker

RHuVλBACK1Fv	5'-GGC GGC TGC GTC AAC ACA GAC TGC GAT CCG CCA CCG CCA GAG-3'
RHuVλBACK2Fv	5'-GCA GGC TGA GTC AGA GCA GAC TGC GAT CCG CCA CCG CCA GAG-3'
RHuVλBACK3aFv	5'-GGT GGC TGA GTC AGC ACA TAG GAC GAT CCG CCA CCG CCA GAG-3'
RHuVλBACK3bFv	5'-GGG TCC TGA GTC AGC TCA GAA GAC GAT CCG CCA CCG CCA GAG-3'
RHuVλBACK4Fv	5'-GGC GGT TGA GTC AGT ATA ACG TGC GAT CCG CCA CCG CCA GAG-3'
RHuVλBACK5Fv	5'-GAC GGC TGA GTC AGC ACA GAC TGC GAT CCG CCA CCG CCA GAG-3'
RHuVλBACK6Fv	5'-TGG GGC TGA GTC AGC ATA AAA TTC GAT CCG CCA CCG CCA GAG-3'

D. *Reamplification with primers containing restriction sites*

Human V_H back primers

HuVH1aBACKSfi	5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG GTG CAG CTG GTG CAG TCT GG-3'
HuVH2aBACKSfi	5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG GTC AAC TTA AGG GAG TCT GG-3'
HuVH3aBACKSfi	5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC GAG GTG CAG CTG GTG GAG TCT GG-3'
HuVH4aBACKSfi	5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG GTG CAG CTG CAG GAG TCG GG-3'
HuVH5aBACKSfi	5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG GTG CAG CTG TTG CAG TCT GC-3'
HuVH6aBACKSfi	5'-GTC CTC GCA ACT GCG GCC CAG CCG GCC ATG GCC CAG GTA CAG CTG CAG CAG TCA GG-3'

Human J_κ forward primers

HuJκ1BACKNot	5'-GAG TCA TTC TCG ACT TGC GGC CGC ACG TTT GAT TTC CAC CTT GGT CCC-3'
HuJκ2BACKNot	5'-GAG TCA TTC TCG ACT TGC GGC CGC ACG TTT GAT CTC CAG CTT GGT CCC-3'
HuJκ3BACKNot	5'-GAG TCA TTC TCG ACT TGC GGC CGC ACG TTT GAT ATC CAC TTT GGT CCC-3'
HuJκ4BACKNot	5'-GAG TCA TTC TCG ACT TGC GGC CGC ACG TTT GAT CTC CAC CTT GGT CCC-3'
HuJκ5BACKNot	5'-GAG TCA TTC TCG ACT TGC GGC CGC ACG TTT AAT CTC CAG TCG TGT CCC-3'

Human J_λ forward primers

HuJλ1FORNOT	5'-GAG TCA TTC TCG ACT TGC GGC CGC ACC TAG GAC GGT GAC CTT GGT CCC-3'
HuJλ2-3FORNOT	5'-GAG TCA TTC TCG ACT TGC GGC CGC ACC TAG GAC GGT CAG CTT GGT CCC-3'
HuJλ4-5FORNOT	5'-GAG TCA TTC TCG ACT TGC GGC CGC ACY TAA AAC GGT GAG CTG GGT CCC-3'

glucose (2×TY-AMP-GLU) and 15% (v/v) glycerol for storage at −70°C as a library stock.

(e) *Rescue of phagemid libraries*

To rescue phagemid particles from the library, 100 ml of 2×TY-AMP-GLU was inoculated with 10^9 bacteria taken from the library stock (approx. 10 μl) and grown for 1·5 h, shaking at 37°C. Cells were spun down (IEC-Centra 8, 4000 revs/min for 15 min) and resuspended in 100 ml of prewarmed (37°C) 2×TY broth containing 100 μg ampicillin/ml (2×TY-AMP), 2×10^{10} plaque-forming units of VCS-M13 (Stratagene) particles were added and the mixture incubated 30 min at 37°C without shaking. The mixture was then added to 900 ml of 2×TY broth containing 100 μg ampicillin/ml and 25 μg kanamycin/ml (2×TY-AMP-KAN), and grown overnight, shaking at 37°C. Phage particles were purified and concentrated by three PEG-precipitations (Sambrook *et al.*, 1990) and resuspended in PBS to 10^{13} transducing units/ml (ampicillin-resistant clones).

(f) *Selection of phOx:BSA binders using tubes*

For selection, 75 mm×12 mm immuno tube (Nunc; Maxisorp) was coated with 4 ml of phOx:BSA (1 mg/ml; 14 phOx per BSA: Mäkelä *et al.*, 1978) in PBS overnight at room temperature. After washing 3 times with PBS, the tube was incubated for 2 h and 37°C with 2% MPBS for blocking. The wash was repeated and phagemid particles (10^{13} t.u.) in 4 ml of 2% MPBS added, incubated 30 min at room temperature, systematically inverting the tube using a rotating turntable, and then left undisturbed for a further 1·5 h at room temperature. Tubes were then washed 20 times with PBS, 0·1% (v/v) Tween 20 and 20 times with PBS (each washing step was performed by pouring buffer in and out immediately). Bound phage particles were eluted from the tube by adding 1 ml of 100 mM-triethylamine, inverting the tube using a rotating turntable for 15 min. The eluted material was immediately neutralized by adding 0·5 ml of 1·0 M-Tris·HCl (pH 7·4). Phage were stored at 4°C. Eluted phage (in 1·5 ml) were used to infect 8 ml of logarithmic growing *E. coli* TG1 cells in 15 ml of 2×TY broth, and plated on TYE-AMP-GLU plates as described above, yielding on average 10^7 t.u. For selection of phOx:BSA binders, the rescue–selection–plating cycle was repeated 4 times, after which phagemid clones were analysed for binding to both phOx:BSA and BSA.

(g) *Selection for lysozyme binders by panning and by affinity column*

A circular Petri dish (35 mm×10 mm Falcon 3001 Tissue culture dish) was used for enrichment by panning. During all steps, the plates were rocked on an A600 rocking plate (Raven Scientific). Plates were coated overnight with 1 ml of TEL (3 mg/ml; Sigma) in 50 mM-sodium hydrogen carbonate (pH 9·6), washed 3 times with 2 ml of PBS, and blocked with 2 ml of 2% MPBS at room temperature for 2 h. Approximately 10^{13} t.u. phage in 1 ml of 2% MPBS were added per plate, and left rocking for 2 h at room temperature. Plates were washed for 5 min with 2 ml of the following solutions: 5 times with PBS; PBS, 0·02% Tween 20; 50 mM-Tris·HCl (pH 7·5), 500 mM-NaCl; 50 mM-Tris·HCl (pH 8·5), 500 mM-NaCl; 50 mM-Tris·HCl (pH 9·5), 500 mM-NaCl and finally 50 mM-sodium hydrogen carbonate (pH 9·6). Bound phage particles were then eluted by adding 1 ml of 100 mM-triethylamine and rocking for 5 min before neutralizing with 0·5 ml of 1 M-Tris·HCl (pH 7·4). Eluted phage was used to infect logarithmic growing *E. coli* TG1 as described above.

Alternatively, TEL-Sepharose columns were used for affinity purification. One ml columns of TEL coupled to Sepharose (as described by Ward *et al.*, 1989) were washed extensively with PBS, blocked with 5 ml of 2% MPBS, and 10^{13} t.u. phage in 1 ml of 2% MPBS loaded. Columns were washed with 50 ml of PBS; 10 ml of PBS, 0·02% Tween 20; 5 ml of 50 mM-Tris·HCl (pH 7·5), 500 mM-NaCl; 5 ml of 50 mM-Tris·HCl (pH 8·5), 500 mM-NaCl; 5 ml of 50 mM-Tris·HCl (pH 9·5) 500 mM-NaCl and finally 5 ml of 50 mM-sodium hydrogen carbonate (pH 9·6), 500 mM-NaCl. Bound phage were eluted using 1·5 ml of 100 mM-triethylamine and neutralized with 0·5 ml 1 M-Tris·HCl (pH 7·4). Eluted phage were used to infect logarithmically growing *E. coli* TG1 as described above.

For selection of lysozyme binders by either method, the rescue–selection–plating cycle was repeated 4 times, after which phagemid clones were analysed for binding by ELISA.

(h) *Rescue of phage or soluble scFv from individual phagemid clones for binding ELISA*

To rescue phage, single ampicillin-resistant colonies, resulting from infection of *E. coli* TG1 with eluted phage, were inoculated into 150 μl of 2 ×TY-AMP-GLU broth in 96-well plates (Cell wells; Corning) and grown with shaking (250 revs/min) overnight at 37°C. A 96-well plate replicator was used to inoculate approximately 4 μl of the overnight cultures on the master plate into 200 μl fresh 2×TY-AMP-GLU. After 1 h, 50 μl of 2×TY-AMP-GLU broth containing 10^8 p.f.u. of VCS-M13 was added to each well, and the plate incubated at 37°C for 45 min without agitation. The plate was then shaken at 37°C for 1 h after which time glucose was removed by spinning down the cells (IEC-Centra 8, 4000 revs/min for 15 min), and aspirating the supernatant with a drawn-out glass Pasteur pipete. Cells were resuspended in 200 μl 2×TY-AMP-KAN broth and grown for 20 h, shaking at 37°C. Supernatant containing phage was tested for binding by ELISA.

To produce soluble scFvs, single ampicillin-resistant colonies of infected *E. coli* HB2151, a non-suppressor strain (Carter *et al.*, 1985), were inoculated into 150 μl of 2×TY broth containing 100 μg ampicillin/ml and 0·1% glucose in 96-well plates and grown with shaking at 37°C until an $A_{600\ nm}$ of 0·9 was reached. Expression of soluble scFv was induced by the addition of isopropyl β-D-thiogalactopyranoside to a final concentration of 1 mM (DeBellis & Schwartz, 1990) and the cultures grown overnight at 30°C. Supernatant containing soluble scFv was taken for analysis by ELISA.

(i) *ELISA*

Analysis of phage for binding to phOx:BSA, BSA or lysozyme by ELISA was performed on bacterial supernatants containing phage essentially as described by Clackson *et al.* (1991), with 100 μg phOx:BSA or BSA/ml, or 3 mg TEL/ml used for coating. The specificity of isolated clones was checked by ELISA of the soluble scFv fragments using plates coated with various proteins at 1 mg/ml (hen egg ovalbumin, hen egg lysozyme, chymotrypsinogen A, cytochrome c, bovine thyroglobin, glyceraldehyde-3-phosphate dehydrogenase, chicken egg white trypsin inhibitor (Sigma), keyhole limpet haemocyanin (CalBiochem)). Binding of soluble scFvs to antigen was detected with the mouse monoclonal antibody 9E10 (1 μg/ml), which recognizes the C-terminal peptide tag (Munro & Pelham, 1986), and peroxidase-conjugated anti-mouse Fc antibody (Sigma), as described (Ward *et al.*, 1989).

(j) *DNA fingerprinting of clones*

The diversity of the original and selected libraries was determined by PCR screening (Güssow & Clackson, 1989). Recombinant clones were screened before and after selection by amplifying the scFv insert using primers LMB3 (5′-CAGGAAACAGCTATGAC, which sits upstream from the pelB leader sequence) and fd-SEQ1 (5′-GAATTTTCT-GTATGAGG, which sits in the 5′ end of gene *3*) followed by digestion with the frequent-cutting enzyme *Bst*NI. The heavy and light chain variable regions from at least 2 clones of each restriction pattern were sequenced using a Sequenase kit (USB) by the dideoxy chain termination method (Sanger *et al.*, 1977). The nucleic acid sequences of the V-regions were compared with a database of germline V-genes to determine the family of origin and extent of somatic mutation.

(k) *Frequency of lambda and kappa light chains in the unselected IgM library*

The frequency of lambda and kappa light chains in the unselected IgM library was determined by probing replica-plated colonies with either an equimolar mixture of the V_λ PCR primers (Table 1) or an equimolar mixture of family-specific V_κ framework 1 probes (Marks *et al.*, 1991). One hundred individual colonies from the unselected IgM library were replica-plated on 2×TY-AMP-GLU plates and lifted onto nylon membranes (Hybond-N, 0·45 μm). The membranes were treated as described (Buluwela *et al.*, 1989) and then ultraviolet crosslinked for 5 min (Stratalinker; Stratagene). Membranes were prehybridized for 20 min at 42°C in hybridization solution (0·9 M-NaCl, 0·09 M-Tris (pH 7·5), 6 mM-EDTA (pH 7·4), 1 mM-sodium pyrophosphate, 0·5% (v/v) NP40, 0·6 mg/l rATP, 20 mg/l yeast RNA, 20 mg/l Ficoll 400, 20 mg/l polyvinylpyrrolidone and 20 mg/l BSA) and then hybridized for 2 h at 42°C with 10 pmol of (γ-^{32}P)-labelled oligonucleotide probe. Membranes were washed once at 42°C for 10 min in 6×SSC (900 mM-NaCl, 90 mM-trisodium citrate, pH 7·0), 0·1% (w/v) SDS, 0·1% (w/v) sodium pyrophosphate, once for 15 min at 55°C in 3 M-tetramethylammonium chloride, 50 mM-Tris (pH 8·0), 0·1% SDS, 2 mM-EDTA and exposed for 2 h on Fuji RX film.

(l) *Purification of scFvs and affinity determination*

The phOx binding scFv clone 15 (αphOx15) and the TEL binding scFv clone 9 (αTEL9), which gave the strongest ELISA signals, were chosen for affinity determination. Colonies of *E. coli* HB2151, a non-suppressor strain, harbouring the appropriate phagemid were used to inoculate 10 l of 2×TY containing 100 μg ampicillin/ml and 0·1% glucose. The cultures were grown to an $A_{600\ nm}$ of 0·9 and expression of soluble scFv induced by the addition of IPTG to a final concentration of 1 mM (DeBellis & Schwartz, 1990). Supernatant was concentrated 8-fold by ultrafiltration (Filtron; Flowgen) and 200 ml loaded onto a 5 ml column of Protein A–Sepharose crosslinked by dimethylpimelidate (Harlow & Lane, 1988) to the monoclonal antibody 9E10 that recognizes the C-terminal peptide tag (Clackson *et al.*, 1991; Munro & Pelham, 1986). The column was washed with 100 ml of PBS; 10 ml of PBS, 0·5 M-NaCl; 10 ml of 0·2 M-glycine (pH 6·0); and

10 ml of 0·2 M-glycine (pH 5·0). The scFv fragment was eluted with 10 ml of 0·2 M-glycine (pH 3·0), neutralized with Tris base and dialysed into PBSE (PBS buffer containing 0·2 mM-EDTA). Supernatant from a separate induction of the αTEL9 scFv was purified on lysozyme–Sepharose (Ward *et al.*, 1989).

Affinities were measured by fluorescence quench techniques, based on the quenching of tryptophan fluorescence by the bound hapten or antigen (Eisen 1964; Foote & Milstein, 1991; J. Foote & G. Winter, unpublished results). All measurements were made with a Perkin–Elmer LS-5B spectrofluorimeter, using an excitation wavelength of 280 nm. Antibody (0·9 ml) in PBSE, was placed in a 4 mm × 10 mm cuvette in the instrument, and held at 20°C.

For determination of the affinity of αphOx15, fluorescence quench titration was performed essentially as described by Foote & Milstein (1991). A regime of hapten excess was used: the antibody concentration (100 nM) was at most equal to the lowest concentration of hapten. Negligible volumes of the hapten 4-γ-amino-butyric acid methylene 2-phenyl-oxazol-5-one (phOx-GABA) were added to αphOx15 protein to cover a concentration range of 0·2 to 4 times the preliminary estimate of the dissociation constant (500 nM), and the fluorescence detemined 1 min after each addition. Emission was monitored at 340 nm. Data were averaged from 3 runs and the value of the equilibrium constant was obtained from a least-squares fit of the data to a hyperbola.

Fluorescence quench titration was also used to determine the affinity of αTEL9 (Eisen, 1964; J. Foote & G. Winter, unpublished results). αTEL9 protein at 200 nM was titrated to 2-fold molar excess with TEL (Sigma) in PBSE, sample fluorescence being determined 1 min after each addition. Emission was monitored at 350 nm and the titration repeated 6 times. Five identical titrations with TEL were also performed on αphOx15 as control. The fluorescence data from each of the 6 titrations of αTEL9 were subtracted from the mean fluorescence values from the 5 control titrations of αphOx15 to account for the fluorescence contributed by the added TEL. To obtain the equilibrium constant, fluorescence data, averaged from the 6 corrected titrations of αTEL9, were fit by least-squares to a hyperbola.

(m) *Western blot*

Western blotting was performed essentially as described by Towbin *et al.* (1979). Samples (10 μg and 1 μg) of TEL were subjected to SDS/PAGE (Laemmli, 1970) and protein transferred by electroblotting to Immobilon-P (Millipore). The blot was blocked with PBS. 3% BSA for 20 min and then incubated with αTEL9 (1 μg/ml) in PBS, 3% BSA for 1·5 h. Binding of αTEL9 to lysozyme was detected with 9E10 (1 μg/ml) and peroxidase-conjugated anti-mouse Fc antibody (Sigma) as described Ward *et al.* (1989).

3. Results

(a) *Generation of scFv gene repertoires and libraries*

Single bands of the correct size for V_H, V_κ and V_λ cDNA were obtained after amplification of first strand cDNA made from RNA primed with the appropriate constant region primer (Table 1). No bands were obtained in the absence of a primer in the first strand cDNA reaction, indicating that the products resulted from the amplification of RNA and not DNA. A major band of the appropriate size for an assembled scFv gene was obtained when the V_H and V_κ, or V_H and V_λ, were combined with linker DNA in a PCR reaction. No product was obtained in the absence of linker DNA (data not shown).

Libraries of $2{\cdot}9 \times 10^7$ $V_{H\mu}$-V_L scFv clones (IgM library) and $1{\cdot}6 \times 10^8$ $V_{H\gamma}$-V_L scFv clones (IgG library) were obtained (Fig. 2). Analysis of 100 colonies from the IgM library by probing revealed that 81 carried either kappa or lambda light chains (45 (56%) for lambda and 36 (44%) for kappa). Analysis of 48 clones from each unselected library (IgM and IgG) indicated that greater than 90% of the clones carried an insert, and the libraries appeared to be extremely diverse as judged by the *Bst*NI restriction pattern (Fig. 3(a)).

(b) *Isolation and characterization of binders*

Phagemid particles were rescued from the library by superinfection with helper phage and selected by passing over either immobilized TEL or phOx:BSA. Eluted phage were used to infect *E. coli*, the library was again rescued with helper phage and the phagemid particles were subjected to a second

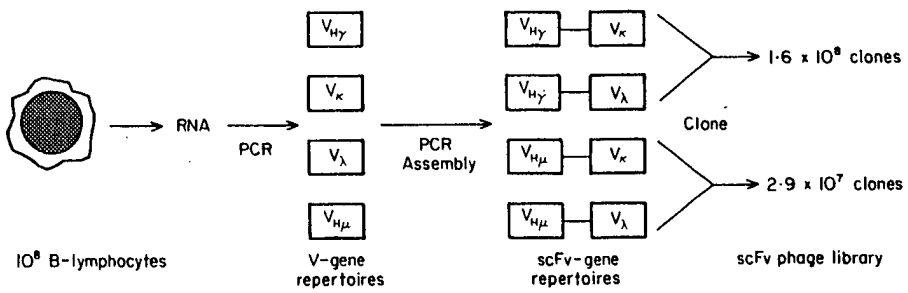


Figure 2. The origin of V-genes in the phage libraries. RNA made from 10^8 B-lymphocytes was primed with constant region-specific primers (for IgM, IgG, Cκ and Cλ) and 1st strand cDNA synthesized. Portions of 1st strand cDNA were used to amplify $V_{H\mu}$ and $V_{H\gamma}$ genes, and V_κ and V_λ genes. The V-genes were assembled together in separate PCR assembly reactions to generate 4 distinct scFv repertoires: $V_{H\mu}$-V_κ, $V_{H\mu}$-V_λ, $V_{H\gamma}$-V_κ and $V_{H\gamma}$-V_λ. The $V_{H\mu}$-V_κ and $V_{H\mu}$-V_λ repertoires were combined and cloned to generate a $V_{H\mu}$ scFv library of $2{\cdot}9 \times 10^7$ clones. Likewise the $V_{H\gamma}$-V_κ and $V_{H\gamma}$-V_λ repertoires were combined and cloned to generate a $V_{H\gamma}$ scFv library of $1{\cdot}6 \times 10^8$ clones.

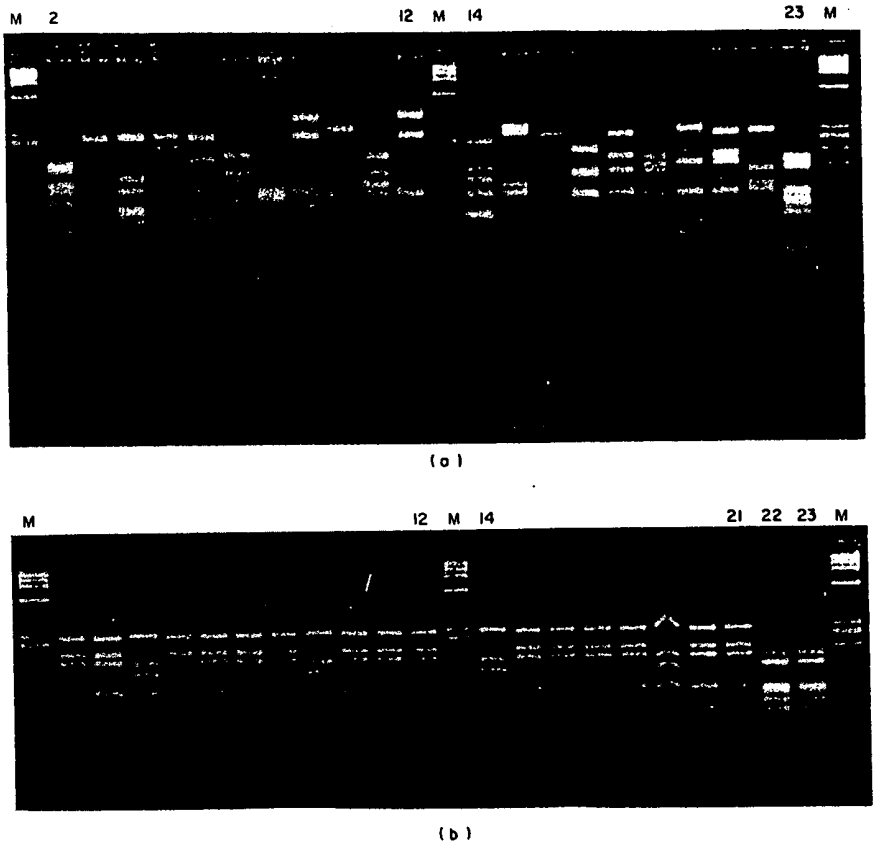


Figure 3. *Bst*N1 fingerprinting of scFv clones. The scFv insert was amplified from individual colonies, the product digested with *Bst*N1 and analysed on an agarose gel. M, ϕX174 DNA *Hae*III-digested molecular weight markers. (a) Lanes 2 to 12 and 14 to 23 are digests from colonies from the library before selection. (b) Lanes 2 to 12 and 14 to 21 are digests from 21 random colonies after 4 rounds of panning of the IgM library on TEL. Lanes 22 and 23 are digests of 2 other TEL binding clones obtained after 4 rounds of selection of the IgM or IgG library on a TEL column, respectively.

round of affinity purification. Four rounds of rescue–selection–infection were performed. Clones binding TEL, BSA and phOx were identified after four rounds of selection of the IgM library (Table 2). In contrast only clones binding TEL were identified after four rounds of selection of the IgG library (Table 2). Unselected clones and clones isolated after one and two rounds of selection showed no binding. Comparison of the frequency of binders to TEL and BSA obtained after three and four rounds of selection indicates up to 50-fold enrichment in the fourth round of selection. Thus, these binders must

Table 2

Frequency of binding clones from scFv libraries before and after selection

	Rounds of selection				
	0	1	2	3	4
A. *IgM library*					
Human anti-TEL: panning	0/864	0/192	0/192	3/192	94/192
Human anti-TEL: columns	—	—	—	—	19/96
Human anti-BSA: panning	0/192	0/192	0/192	2/192	43/96
Human anti-phOx: panning	0/192	0/192	0/192	0/192	1/96
B. *IgG library*					
Human anti-TEL: panning	—	—	—	—	0/96
Human anti-TEL: columns	—	—	—	—	6/96
Human anti-BSA: panning	—	—	—	—	0/96
Human anti-phOx: panning	—	—	—	—	0/96

Panning, antigen coated on Petri dish; columns, antigen covalently linked to Sepharose column; IgM library, single chain Fv library (scFv) with V_H genes derived from IgM mRNA; IgG library, scFv genes with V_H genes derived from IgG mRNA.

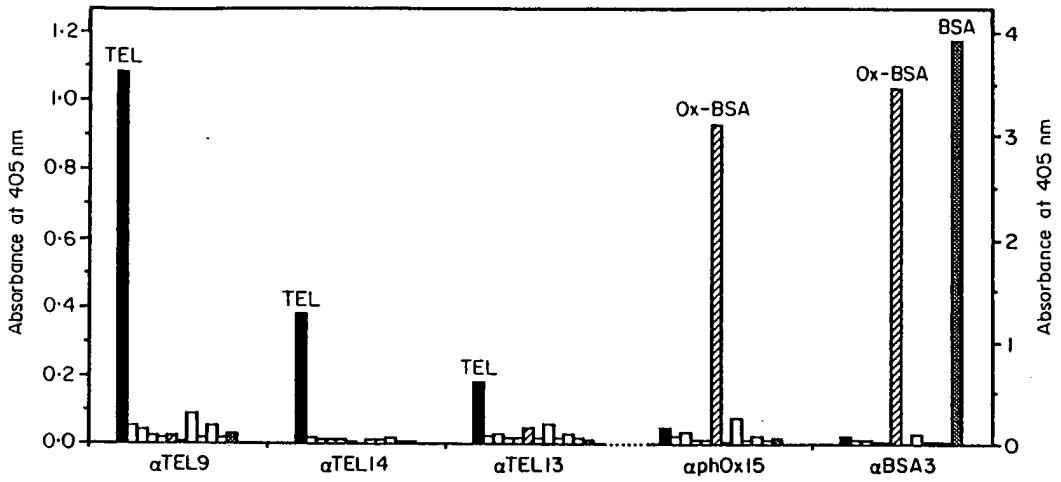


Figure 4. Specificity of soluble single chain Fvs (scFvs). Binding was determined by ELISA to a variety of proteins. αTEL9, αTEL13 and αTEL14 = 3 anti-turkey lysozyme scFvs; αphOx15 = anti-2-phenyloxazole-5-one scFv; αBSA3 = anti-bovine serum albumin scFv. Antigens: TEL (filled box), phOx-BSA (hatched box), BSA (stippled box); other antigens (open box) = keyhole limpet haemocyanin, bovine thyroglobulin, chymotrypsinogen A, hen-egg ovalbumin, cytochrome *c*, hen egg lysozyme, hen egg trypsin inhibitor, glyceraldehyde-3-phosphate dehydrogenase, and plastic. plastic.

have been present in the original library at a frequency of 1 per $6{\cdot}25 \times 10^6$ clones ($1/50^4$) if enrichment were equal over the four rounds of selection.

*Bst*NI fingerprinting of 23 lysozyme binding clones from the IgM library indicated the presence of three different digestion patterns, whereas the six lysozyme binding clones obtained from the IgG library all had the same restriction pattern (Fig. 3(b), and data not shown). The *Bst*NI fingerprinting of 35 BSA binding clones indicated the presence of only one digestion pattern (data not shown) which was different from the pattern of the phOx binding clone.

The sequences of the variable regions of multiple clones representing the different restriction patterns indicated that there were four unique TEL binders (αTEL9, αTEL13, αTEL14 and αTEL16), one BSA binder (αBSA3) and one phOx binder (αphOx15) (Table 3). The V_Hs were derived from four different V_H families and five different V_H germline genes (Table 5). The light chains were mainly lambda (5/6) and were derived from four different light chain families and germline genes (Table 5). Both V-genes of αBSA3 were unmutated compared to germline (Tables 4 and 5). Similarly, the V-genes of αphOx15 were minimally mutated from germline (4 differences with VH380·6 (Berman *et al.*, 1988) and six with IGLV3S1 (Frippiat *et al.*, 1990)). Two other antibodies (αTEL13 and αTEL16) had heavy chains that are more extensively mutated (11 and 18 changes from VH251 (Sanz *et al.*, 1989)). Only upper estimates of mutation are possible for the other chains (Tables 4 and 5), as the sequences of all the germ-line V-genes from these families are not known. Finally, the TEL binder isolated from the IgG library (αTEL16) was highly related to one of the IgM TEL binders (αTEL13), and with a greater degree of somatic mutation.

(c) *Specificity of binding*

Soluble antibody fragments were readily prepared by growth of *E. coli* HB2151, a non-suppressor strain, carrying the phagemid (Hoogenboom *et al.*, 1991). Soluble scFvs of αphOx15, αBSA3, αTEL9, αTEL13 and αTEL14 were highly specific in an ELISA to test cross-reactivity (Fig. 4). The αTEL16 scFv, isolated from the IgG library, could not be detected in ELISA as a soluble fragment, probably due to its low affinity.

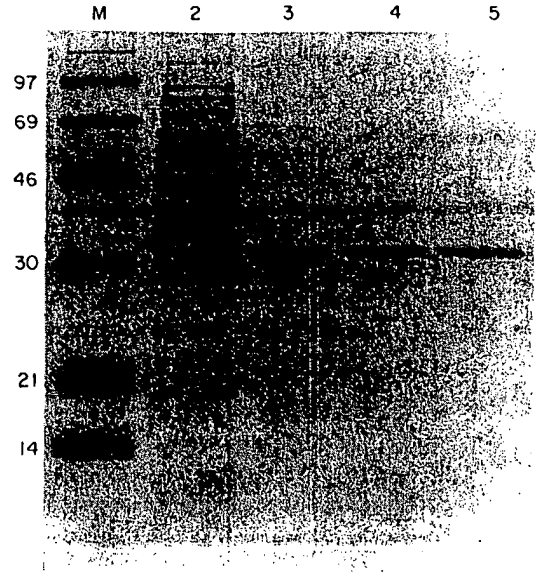


Figure 5. Purification of scFvs protein from a bacterial supernatant. M, molecular weight markers ($\times 10^{-3}$). Lane 2, unpurified bacterial supernatant; lane 3, αTEL9 scFv protein purified on a lysozyme–Sepharose column; lane 4, αTEL9 scFv protein purified on column of antibody 9E10 directed against the c-*myc* tag; lane 5, αphOx15 scFv protein purified as in lane 4.

Table 3

Deduced protein sequences of antigen-specific heavy and light chains selected from unimmunized libraries

A. *Heavy chains*

Clone	FR 1	CDR 1	FR 2	CDR 2	FR 3	CDR 3	FR 4
αphOx15	QVQLVQSGAEVKKPGASVKVSCKASGYTFT	SYGIS	WVRQAPGQGLEWMG	WISAYNGNTKYAQKLQG	RVTMTTDTSTSTAYMELRSLRSDDTAVYYCVR	LLPKRTATLHYYIDV	WGKGTLVTVSS
αBSA3	QVQLVQSGGGVVQPGRSLRLSCAASGFTFS	SYGMH	WVRQAPGKGLEWVA	VISYDGSNKYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK	TGYSSGWGYFDY	WGQGTLVTVSS
αTEL9	QVQLQQSGSGLVKPSQTLSLTCSVSGDSIS	SGGYSWS	WIRQPSGKGLEWIG	SVHHSGPTYYNPSLKS	RVTMSVDTSKNQFSLKLKSVTAADTAMYFCAR	EGGSTWRSLYKHYYMDV	WGKGTLVTVSS
αTEL14	QVQLQESGPGLVKPSETLSLVCTVSGGSLS	FSYWG	WIRQPPGKGLEWIG	YISHRGTDYNSSLQS	RVTISADTSKNQFSLKLSSVTAADTAVYYCAR	SFSNSFFFGY	WGQGTLVTVSS
αTEL13	QVQLVQSGAEVKKPGQSLMISCQGSGYSFS	NYWIG	WVRQMPGKGLEWMG	IIYPGDSDTRYSPSFQG	QVTISADKSISTAYLHWSSLKASDTALYYCAR	LVGGTPAY	WGQGTLVTVSS
αTEL16	QVQLVQSGAEVKKPGQSLRISCKGAGYSFS	TYWIG	WVRQMPGKGLEWMG	IIYPDDSDTRYSPSFEG	QVTISVDKSITTAYLHWSSLKASDTAIYYCAR	LVGGAPAY	WGQGTLVTVSS

B. *Light chains*

Clone	FR 1	CDR 1	FR 2	CDR 2	FR 3	CDR 3	FR 4
αphOx15	QSVLTQPPSVSAAPGQKVTISC	SGSSSNIGNNYVS	WYQHLPGTAPNLLIY	DNNKRPS	GIPDRFSGSKSGTSATLGITGLQTGDEADYYC	GTWDGRLTAAV	FGSGTKVTVLG
αBSA3	SSELTQDPAVSVALGQTVRITC	QGDSLRSYYAS	WYQQKPGQAPVLVIY	GKNNRPS	GIPDRFSGSSSGNTASLTITGAQAEDEADYYC	NSRDSSGNHVV	FGGGTKLTVLG
αTEL9	EIVLTQSPSSLSASVGDRVTITC	RASQSISNYLN	WYQQKPGKAPKLLIY	AASTLQS	GVPSRFSGSGSGTDFTLTINSLQPEDFATYYC	QQTNSFPLT	FGGGTKLEIKR
αTEL14	SSELTQDPAVSVAFGQTVRITC	QGDSLRSSYAS	WYQQKPGQAPLLVIY	GENSRPS	GIPDRFSGSSSGNTASLTITGAQAEDEADYYC	NSRDSRGTHLEV	FGGGTKLTVLG
αTEL13	HVILTQPASVSGSPGQSITISC	TGSSRDVGGYNYVS	WYQHHPGKAPKLLIS	EVTNRPS	GVSNRFSGSKSGNTASLTISGLQAEDEADYFC	ASYTSSKTYV	FGRGTKLTVLG
αTEL16	QSALTQPASVSGSPGQSITISC	SGSSSDIGRYDYVS	WYQHYPDKAPKLLIY	EVKHRPS	GISHRFSASKSGNTASLTISELQPGDEADYYC	ASYTESKTYI	FGGGTKVTVLG

FR, framework region; CDR, complementarity-determining region.

Table 4

Nucleotide sequences of antigen specific heavy and light chain V-genes selected from unimmunized libraries compared with the most homologous germline gene

A. Heavy chains

```
                  10        20        30        40        50        60        70        80        90       100
VH380.6   CAGGTTCAGC TGGTGCAGTC TGGAGCTGAG GTGAAGAAGC CTGGGGCCTC AGTGAAGGTC TCCTGCAAGG CTTCTGGTTA CACCTTTACC AGCTATGGTA
αphOx15   -----g---- ---------- ---G------ ---------- ---------- ---------- ---------- ---------- ---------- ----------
                 110       120       130       140       150       160       170       180       190       200
VH380.6   TCAGCTGGGT GCGACAGGCC CCTGGACAAG GGCTTGAGTG GATGGGATGG ATCAGCGCTT ACAATGGTAA CACAAACTAT GCACAGAAGC TCCAGGGCAG
αphOx15   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------G--- ---------- ----------
                 210       220       230       240       250       260       270       280       290
VH380.6   AGTCACCATG ACCACAGACA CATCCACGAG CACAGCCTAC ATGGAGCTGA GGAGCCTGAG ATCTGACGAC ACGGCCGTGT ATTACTGTGC GAGA
αpHOx15   ---------- ---------- ---------- ---------- -----A---- ---------- ---------- ---------- ----------T ----

                  10        20        30        40        50        60        70        80        90       100
VH1.9III  CAGGTGCAGC TGGTGGAGTC TGGGGGAGGC GTGGTCCAGC CTGGGAGGTC CCTGAGACTC TCCTGTGCAG CCTCTGGATT CACCTTCAGT AGCTATGGCA
αBSA3     ---------- -----c---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
                 110       120       130       140       150       160       170       180       190       200
VH1.9III  TGCACTGGGT CCGCCAGGCT CCAGGCAAGG GGCTGGAGTG GGTGGCAGTT ATATCATATG ATGGAAGTAA TAAATACTAT GCAGACTCCG TGAAGGGCCG
αBSA3     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
                 210       220       230       240       250       260       270       280       290
VH1.9III  ATTCACCATC TCCAGAGACA ATTCCAAGAA CACGCTGTAT CTGCAAATGA ACAGCCTGAG AGCTGAGGAC ACGGCTGTGT ATTACTGTGC GAAA
αBSA3     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----

                  10        20        30        40        50        60        70        80        90       100
U514A     CAGGTGCAGC TGCAGGAGTC GGGCCCAGGA CTGGTGAAGC CTTCACAGAC CCTGTCCCTC ACCTGCACTG TCTCTGGTGG CTCCATCAGC AGTGGTGGTT
αTEL9†    -----a---- -----c---- a---T----- ---------- ---------- ---------- ------T--- ----------A -------TC- ----------
U514G     ---------- ---------- ---------- ---------- ----GG---- ---------- ------G--- --------TA ---------- ----------
                 110       120       130       140       150       160       170       180       190       200
U514A     ACTACTGGAG CTGGATCCGG CAGCCCCCAG GGAAGGGACT GGAGTGGATT GGGTACATCT ATTACAGTGG GAGCACCTAC TACAACCCGT CCCTCAAGAG
αTEL9     ---C------ ---------- -----AT--- -------G-- ---------- ---AGTG--C --C------- -CC------- ---------- ----------
U514G     --------G- ---------- ---------- -------G-- ---------- ---AGT----- --C-T----- ---------- ---------- ----------
                 210       220       230       240       250       260       270       280       290
U514A     TCGAGTTACC ATATCAGTAG ACACGTCTAA GAACCAGTTC TCCCTGAAGC TGAGCTCTGT GACTGCCGCG GACACGGCCG TGTATTACTG TGCGAGA
αTEL9     ------C--- --G-----T- -------C-- ---------- --------AT ---AG----- ---C-----A ----------A ------TT-- -------
U514G     ------C--- ---------- -------C-- ---------- ---------- ---------- ---C-----A ---------- ---------- -------

                  10        20        30        40        50        60        70        80        90       100
U4H       CAGGTGCAGC TGCAGGAGTC GGGCCCAGGA CTGGTGAAGC CTTCGGAGAC CCTGTCCCTC ACCTGCACTG TCTCTGGTGG CTCCATCAGT AGTTACTACT
αTEL14    ---------- ---------- ---G------ ---------- ---------- ---------- GT-------- ---------- ----C----- TT--C-----
                 110       120       130       140       150       160       170       180       190       200
U4.H      GGAGCTGGAT CCGGCAGCCC CCAGGGAAGG GACTGGAGTG GATTGGGTAT ATCTATTACA GTGGGAGCAC CAACTACAAC CCCTCCCTCA AGAGTCGAGT
αTEL14    --G------- ---------- ---------- ---------- ------C--C ----C-C--- -G---  -- -G-------T T--------C ----------
                 210       220       230       240       250       260       270       280       290
U4.H      CACCATATCA GTAGACACGT CCAAGAACCA GTTCTCCCTG AAGCTGAGCT CTGTGACCGC TGCGGACACG GCCGTGTATT ACTGTGCGAG A
αTEL14    ---------- -C-------- ---------- ---------- ---T------ ---------- ---------- ---------- ---------- -
```

Table 4 *(continued)*

```
                    10        20        30        40        50        60        70        80        90       100
VH251       GAGGTGCAGC TGGTGCAGTC TGGAGCAGAG GTGAAAAAGC CCGGGGAGTC TCTGAAGATC TCCTGTAAGG GTTCTGGATA CAGCTTTACC AGCTACTGGA
αTEL13      c--------- ---------- ---------- ---------- -----C---- -----T---- ------C--- ---------- -------G- -A-------
αTEL16      c--------- ---------- ---G------ ---------- -----C---- -----G---- ---------- --G------- -------G- -C-------
                   110       120       130       140       150       160       170       180       190       200
VH251       TCGGCTGGGT GCGCCAGATG CCCGGGAAAG GCCTGGAGTG GATGGGGATC ATCTATCCTG GTGACTCTGA TACCAGATAC AGCCCGTCCT TCCAAGGCCA
αTEL13      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
αTEL16      ---------- ---------- ---------- ---------- ---------- ---------- A--------- ---------- --T------- --G-G-----
                   210       220       230       240       250       260       270       280       290
VH251       GGTCACCATC TCAGCCGACA AGTCCATCAG CACCGCCTAC CTGCAGTGGA GCAGCCTGAA GGCCTCGGAC ACCGCCATGT ATTACTGTGC GAGA
αTEL13      ---------- ---------- ---------- ---------- --C--C---- ---------- ---------- -------C-T- -C-------- ---G
αTEL16      ---------- --C-T----- ---------C T-----A--- -----C---- ---------- ---------- --------T- ---------- ----
```

B. *Light chains*

```
                    10        20        30        40        50        60        70        80        90       100
JMA1A       CAGTCTGTGT TGACGCAGCC GCCCTCAGTG TCTGCGGCCC CAGGACAGAA GGTCACCATC TCCTGCTCTG GAAGCAGCTC CAACATTGGG AATAATTATG
αph0x15     ---------- ---------- ---------- ---------- --------A- ---------- ---------- ---------- ---------- ----------
                   110       120       130       140       150       160       170       180       190       200
JMA1A       TATCCTGGTA CCAGCAGCTC CCAGGAACAG CCCCCAAACT CCTCATTTAT GACAATAATA AGCGACCCTC AGGGATTCCT GACCGATTCT CTGGCTCCAA
αph0x15     ---------- ------C--- ---------- -------T-- ---------- ---------- ---------- ---------- ---------- ----------
                   210       220       230       240       250       260       270       280       290
JMA1A       GTCTGGCACG TCAGCCACCC TGGGCATCAC CGGACTCCAG ACTGGGGACG AGGCCGATTA TTACTGCGGA ACATGGGATA GCAGCCTGAG TGCT
αph0x15     ---------- ---------- ---------- ---------- ---------- ---------- ---------- --C------G --C------C ----
```

```
                    10        20        30        40        50        60        70        80        90       100
IgLV3S1     TCTTCTGAGC TGACTCAGGA CCCTGCTGTG TCTGTGGCCT TGGGACAGAC AGTCAGGATC ACATGCCAAG GAGACAGCCT CAGAAGCTAT TATGCAAGCT
αBSA3       --g------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
αTEL14      --g------- ---------- ---------- ---------- -C-------- ---------- ---------- -------T-- ------T-C- ----------
                   110       120       130       140       150       160       170       180       190       200
IgLV3S1     GGTACCAGCA GAAGCCAGGA CAGGCCCCTG TACTTGTCAT CTATGGTAAA AACAACCGGC CCTCAGGGAT CCCAGACCGA TTCTCTGGCT CCAGCTCAGG
αBSA3       ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
αTEL14      ---------- ---------- ---------C ---------- -------G-- ----G----- ---------- ---------- ---------- ----------
                   210       220       230       240       250       260       270       280
IgLV3S1     AAACACAGCT TCCTTGACCA TCACTGGGGC TCAGGCGGAA GATGAGGCTG ACTATTACTG TAACTCCCGG GACAGCAGTG GTAACCAT
BSA3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- --------
αTEL14      ---------- ---------- ---------- ---------- -----A---- ---------- ---------- --------G- ----C-----
```

	10	20	30	40	50	60	70	80	90	100
JMVλ2F.1	CAGTCTGCCC	TGACTCAGCC	TGCCTCCGTG	TCTGGGTCTC	CTGGACAGTC	GATCACCATC	TCCTGCACTG	GAACCAGCAG	TGATGTTGGG	AGTTATAACC
αTEL13	--cgt-ata-	-------a--	----------	----------	----------	----------	----------	---G------	A--C-----T	G---------T
αTEL16	--------t-	----------	----------	----------	----------	----------	-------G--	---G------	---CA----T	C-----G--T
	110	120	130	140	150	160	170	180	190	200
JMVλ2F.1	TTGTCTCCTG	GTACCAACAG	CACCCAGGCA	AAGCCCCCAA	ACTCATGATT	TATGAGGGCA	GTAAGCGGCC	CTCAGGGGTT	TCTAATCGCT	TCTCTGGCTC
αTEL13	A---------	---T--G--T	----------	----------	----C-A---	-C-----T--	C---T-----	G---------	--C-------	----------
αTEL16	A---------	---------T	T----T-A--	----------	------C---	-------T--	AAC-T-----	-------A-A	--C-------	-------C--
	210	220	230	240	250	260	270	280	290	
JMVλ2F.1	CAAGTCTGGC	AACACGGCCT	CCCTGACAAT	CTCTGGGCTC	CAGGCTGAGG	ACGAGGCTGA	TTATTACTGC	AGCTTATATA	CAAGCAGCAG	CACT
αTEL13	A---------	----------	-------C--	T---------	---------A	----------	-----T----	GC--C-----	----TTC--A	G---
αTEL16	------C---	----------	-------C--	-----A----	---C---GA-	----------	----------	GC--C-----	--GAA--T-A	G---

	10	20	30	40	50	60	70	80	90	100
HK137	GACATCCAGA	TGACCCAGTC	TCCATCCTCA	CTGTCTGCAT	CTGTAGGAGA	CAGAGTCACC	ATCACTTGTC	GGGCGAGTCA	GGGCATTAGC	AATTATTTAG
αTEL9	g-a--tgtgt	----g-----	---------C	----------	----------	----------	---------C-	-----A-----	-A--------	--C-------A
	110	120	130	140	150	160	170	180	190	200
HK137	CCTGGTTTCA	GCAGAAACCA	GGGAAAGCCC	CTAAGCTCCT	GATCTATGCT	GCATCCAGTT	TGCAAAGTGG	GGTCCCATCA	AGGTTCAGCG	GCAGTGGATC
αTEL9	AT----A---	----------	----------	----------	----------	-------C--	----------	----------	----------	----------
	210	220	230	240	250	260	270	280		
HK137	TGGGACAGAT	TTCACTCTCA	CCATCAGCAG	CCTGCAGCCT	GAAGATTTTG	CAACTTATTA	CTGCCAACAG	TATAATAGTT	ACCCT	
αTEL9	----------	----------	------A---	----------	----------	-------C--	T--T------	AC---C----	TT--G	

Lower case, differences from germline genes encoded by the PCR primer, complementarity-determining regions (CDRs) are underlined. αBSA3, bovine serum albumin binder; αphOx15, 2-phenyloxazol-5-one binder; αTEL9, αTEL13, αTEL14 and αTEL16, turkey egg lysozyme binders. References for germline genes: VH380·6, U514A, U514G, U4H, JMλ1A and JMλ2F: M. B. Llewelyn, J. D. Marks, I. M. Tomlinson, G. Walter & G. Winter, unpublished results; VH1·9III: Berman *et al.* (1988); VH251: Sanz *et al.* (1989); IgLV3S1: Frippiat *et al.* (1990); HK137: Bentley & Rabbitts (1983). Nucleotide and protein sequences have been deposited with the European Molecular Biology Library (accession numbers X61640 to X61651 inclusive). phOx-binding phage with a mouse heavy chain and human light chains were identified in addition to the entirely human αphOx15, but are not included in this paper. The mouse heavy chain corresponded to the VHB domain of Clackson *et al.* (1991) that had been isolated in the same laboratory, and presumably arose from contamination during the library construction. This demonstrates the importance of completely sequencing all antibody constructs.

† αTEL9 appears to be derived partially from germline genes U514A and U514G, suggesting that it is a result of PCR cross-over between 2 highly related V_Hs.

Table 5

V-gene family, germline derivation and extent of somatic hypermutation of antigen-specific clones isolated from unimmunized libraries

	V_H			V_L		
Clone	Family	Germline gene	Differences from germline	Family	Germline gene	Differences from germline
αBSA3	V_{H3}	VH1-9III	0	$V_{\lambda 3}$	IGLV3S1	0
αphOx15	V_{H1}	VH380-6	4	$V_{\lambda 1}$	JMVλ1A	7
αTEL9	V_{H4}	U514A (U514G)	<22	$V_{\kappa 1}$	HK137	<20
αTEL14	V_{H4}	U4-H	<19	$V_{\lambda 3}$	IGLV3S1	<10
αTEL13	V_{H5}	VH251	11	$V_{\lambda 2}$	JMVλ2F	<31
αTEL16	V_{H5}	VH251	18	$V_{\lambda 2}$	JMVλ2F	<38

αBSA3, bovine serum albumin binder; αphOx15, 2-phenyl-oxazol-5-one binder; αTEL9, αTEL13, αTEL14 and αTEL16, turkey egg lysozyme binders. References for germline genes: see Table 4.

(d) *Protein purification and binding affinity*

Soluble scFv αTEL9 was purified in one step on a TEL-Sepharose column or *via* its c-*myc* peptide tag on a 9E10 antibody column (Fig. 5). Soluble scFv αphOx15 was purified in one step on a 9E10 column (Fig. 5). Typical yields were 2 mg/l after purification on 9E10 and 5 to 10 mg/l after purification on an antigen column. The dissociation constant of the αTEL9 scFv was 86(±61) nM and the dissociation constant of the αphOx15 scFv was 534(±72) nM. The high standard error observed for the dissociation constant of αTEL9 has been observed for hen egg lysozyme binding antibodies using this technique. However, equilibrium constants obtained by fluorescence quench titration are consistent with those deduced by the more precise pseudo-equilibrium relaxation method (J. Foote & G. Winter, unpublished results). Finally, soluble αTEL9 scFv could be used to detect lysozyme (1 μg) in a Western blot (data not shown).

4. Discussion

We used a phage display library utilizing V-gene repertoires to isolate antibody fragments of reasonable affinity against three different (foreign) antigens. The two donors were unimmunized, and their serum IgM antibodies did not appear to bind to the antigens TEL or phOx-BSA as there was no difference in signal intensity in wells coated with antigen compared with control wells not containing antigen. Furthermore most of the VH genes of the binders derive from the IgM (naive and primary response B-cells) rather than the IgG mRNA (secondary response B-cells). Each of the heavy and light chain pairings in Table 5 is unique and contrasts with the promiscuous pairings (in which one chain is associated with more than one partner) noted in libraries from the IgG mRNA from immunized animals (Clackson *et al.*, 1991; Caton & Koprowski, 1990; Persson *et al.*, 1991). Therefore, the library appears to be naive with respect to these antigens.

A recent attempt to isolate human antibodies from an unimmunized donor using a λ phage random combinatorial library failed (Persson *et al.*, 1991). The library (10^6 members) was constructed from IgG mRNA using only PCR primers for V_{H1}, V_{H3}, $V_{\kappa 1}$ and $V_{\kappa 3}$ gene families and was screened for antigen binding using nitrocellulose filters. However, library size, diversity and binding threshold determine the chances of isolating binders. The probability (p) that an epitope is not recognized by at least one antibody in a library depends on the probability ($p[K]$) that an individual antibody recognizes a random epitope with an affinity above a threshold value ($[K]$) and on the number of different antibodies (N) according to the equation $p = e^{-Np[K]}$ (Perelson, 1989).

We attempted to maximize the size of the library by using a pUC-based phagemid (Hoogenboom *et al.*, 1991) that has higher transformation efficiencies than fd vectors. Indeed our library sizes (10^7 to 10^8 members) were at least an order of magnitude greater than with phage fd (Clackson *et al.*, 1991). We also attempted to maximize diversity by using primers optimized for each V-gene family, as well as utilizing IgG and IgM mRNA and both κ and λ light chains. The V_H genes of the binders belong to four different families (V_H families 1, 3, 4 and 5), as do the light chain genes (V_λ families 1, 2 and 3, and V_κ family 1). Furthermore, most (5/6) of the binders were derived from the IgM mRNA, perhaps reflecting the greater diversity of V_H genes. Indeed the only binder from the IgG mRNA (αTEL16) had the poorest binding affinity and/or decreased expression and was barely detectable by ELISA.

The chances of finding a phage with binding activity also depend on its affinity and the efficiency and number of rounds of selection. Both phage (McCafferty *et al.*, 1990; Scott & Smith, 1990) and phagemid (Hoogenboom *et al.*, 1991; Bass *et al.*, 1990) vectors have been used to display peptide or protein fusions with g3p. The phage vectors allow three copies of the g3p fusion protein on each phage particle (Glaser-Wuttke *et al.*, 1989), whereas the g3p fusion protein encoded by phagemid vectors has to compete with the g3p of the helper phage for incorporation into the phagemid particle. Although phage vectors should permit isolation of a greater

number of binders, by virtue of the avidity of binding of the multivalent antibody heads, many will have poor affinities. To enrich for the higher-affinity antibodies, we used phagemid vectors. We noted lower selection efficiencies with phagemid (50-fold/round), compared to 675 to 1000-fold per round for phage vectors (Clackson *et al.*, 1991; McCafferty *et al.*, 1990). We found that three or four rounds of selection were required to isolate the binders, and estimate that only one or two copies of each were present in the original library of 3×10^7 members.

The binders utilize both germline and mutated V-genes. Most of the differences are likely to have arisen as a result of somatic mutation of the V-genes in the original B-cells, but some may have arisen during the PCR amplification and assembly process. Indeed the heavy chain of αTEL9 may have arisen from a cross-over during PCR amplification between rearranged V_H-genes from two highly related germline genes U514A and U514G (Table 4). Surprisingly, most of the binders (5/6) utilized V_λ rather than V_κ genes despite their equal representation in the unselected library. However, human hybridomas prepared by EBV immortalization often secrete IgM and λ chains (Thompson *et al.*, 1991), and during maturation of the immune response, the repertoire may shift from IgM, λ antibodies to IgG, κ (Thompson *et al.*, 1991; J. Bye, N. Hughes-Jones, J. D. Marks & G. Winter, unpublished results).

By using phagemid vectors we can mimic the switch of antibody from its display on B-cells to its secretion by plasma cells. By interposing a stop codon between the antibody and g3p, the antibody fragments can be switched between surface display, or secretion as a soluble fragment from bacteria, by growth in suppressor or non-suppressor strains of bacteria (Hoogenboom *et al.*, 1991). The affinities of two of the soluble antibody fragments αphOx15 and αTEL9, prepared in this way, as determined by fluorescence quench ($K_a = 2 \times 10^6$ M^{-1} and 10^7 M^{-1}, respectively), appear similar to those of human IgM antibodies derived from PBLs after immunization. For example the affinities of human IgM antibodies directed against rhesus D antigen, and made by EBV immortalization of PBLs from immunized donors lie in the range of 10^7 M^{-1} (Hughes-Jones & Gorick, 1991).

The antibody fragments isolated from the library are also highly specific (Fig. 4) to the antigen used in panning. For example, those fragments isolated using TEL did not bind to a range of other protein antigens, including hen egg white lysozyme that differs by only seven amino acids (Imoto *et al.*, 1972). The monovalent αTEL9 fragment could even be used in Western blotting but the sensitivity (1 μg TEL) was poor.

Although we can make human antibodies with reasonable affinity and specificity, a yet more diverse and large library should enable the isolation of even higher-affinity antibodies (Perelson, 1989). For example, the rearranged V_H genes would reflect more the naive B-cell repertoire if they had been prepared from the mRNA of membrane-bound IgM or IgD (for example, by basing primers for cDNA synthesis in the membrane anchor region). Other diverse libraries might be constructed by assembling unrearranged V-genes with synthetic D and J elements, or by assembling diverse antigen binding loops on a common structural framework (Milstein, 1990). Larger libraries could be made by improving transfection and ligation efficiencies and by scale-up, or by encoding repertoires of light chains on one vector and heavy chains on another (Hoogenboom *et al.*, 1991).

Alternatively higher-affinity antibodies might be made by mutating the binders and selecting those with improved affinity (Winter & Milstein, 1991). Point mutants could be made in a variety of ways; for example, using an error-prone polymerase (Liao & Wise, 1990), spiked oligonucleotides (Hermes *et al.*, 1989), or growth of the phage in mutator strains of bacteria (Schaaper, 1988; Yamagishi *et al.*, 1990). For more extensive variation, artificial cross-overs could be induced with related genes using the polymerase chain reaction (Meyerhans *et al.*, 1990), or light or heavy chains replaced by repertoires (Clackson *et al.*, 1991). Selection of antibodies on phage according to affinity has demonstrated that, for example, high-affinity binding phage (10^8 M^{-1}) can be fractionated 10^4-fold with respect to low-affinity phage (10^5 M^{-1}) using only two rounds of selection (Clackson *et al.*, 1991). By using several rounds of selection and adjusting the coating density of the antigen used for panning, it is also possible to select between phages bearing antibodies that are much closer in affinity. However, phagemid vectors leading to display of only a single copy of the antibody on the surface of the phage are preferable for selection between phages with closely related affinities when using antigen immobilized on solid phase (T.P.B. & G.W., unpublished results).

For making high-affinity antibodies, phage display libraries built from the spleen mRNA of hyperimmunized animals (Clackson *et al.*, 1991), or PBL mRNA of deliberately immunized humans remain attractive. However, immunization is often difficult, and new libraries have to be constructed for each antigen. In contrast, a single library made without immunization may provide a rich source of antibody specificities, including those directed against "naive" antigens (as described above), common pathogens or self antigens. For example, from the same library as above, we have isolated specificities directed to human blood group B, human tumour necrosis factor-α, and a human monoclonal antibody (our unpublished results). We propose the term "natural" libraries for those derived from unimmunized donors, and envisage that human antibodies of many specificities will be made in the future by panning a single large natural phage display library with antigen.

We thank R. Pannell for 9E10 antibody, M. Hobart for the sheep anti-M13 antibody, and W. Ouwehand and

C. Milstein for advice and encouragement. J.D.M. was supported by the Medical Research Council AIDS directed program, H.R.H. by the D. Collen Research Foundation, Leuven, and the European Molecular Biology Organization, and A.D.G. by the Cancer Research Campaign. T.P.B. was the recipient of a Medical Research Council studentship.

References

Bass, S., Greene, R. & Wells, J. A. (1990). Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins*, **8**, 309–314.

Bentley, D. L. & Rabbitts, T. H. (1983). Evolution of immunoglobulin V genes: evidence indicating that recently duplicated human V sequences have diverged by gene conversion. *Cell*, **32**, 181–189.

Berman, J. E., Mellis, S. J., Pollock, R., Smith, C. L., Suh, H., Heinke, B., Kowal, C., Surti, U., Cantor, C. R. & & Alt, F. W. (1988). Content and organization of the human IgVH locus: Definition of three new VH families and linkage to the Ig CH locus *EMBO J.* **7**, 727–738.

Buluwela, L., Foster, A., Boehm, T. & Rabbitts, T. H. (1989). A rapid procedure for colony screening using nylon filters. *Nucl. Acids Res.* **17**, 452.

Carter, P., Bedouelle, H. & Winter, G. (1985). Improved oligonucleotide site-directed mutagenesis using M13 vectors. *Nucl. Acids Res.* **13**, 4431–4443.

Cathala, G., Savouret, J., Mendez, B., West, B. L., Karin, M., Martial, J. A. & Baxter, J. D. (1983). A method for isolation of intact, transcriptionally active ribonucleic acid. *DNA*, **2**, 329–335.

Caton, A. J. & Koprowski, H. (1990). Influenza virus hemagglutinin-specific antibodies isolated from a combinatorial expression library are closely related to the immune response of the donor. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 6450–6454.

Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature (London)*, **352**, 624–628.

DeBellis, D. & Schwartz, I. (1990). Regulated expression of foreign genes fused to *lac*: control by glucose levels in growth medium. *Nucl. Acids. Res.* **18**, 1311.

Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucl. Acids Res.* **16**, 6127–6145.

Eisen, H. N. (1964). Determination of antibody affinity for haptens and antigen by means of fluorescence quenching. *Methods Med. Res.* **10**, 115–121.

Foote, J. & Milstein, C. (1991). Kinetic maturation of an immune response. *Nature (London)*, **352**, 530–532.

Frippiat, J. P., Chuchana, P., Bernard, F., Buluwela, L., Lefranc, G. & Lefranc, M. P. (1990). First genomic sequence of a human Ig variable lambda gene belonging to subgroup III. *Nucl. Acids Res.* **18**, 7134.

Gibson, T. J. (1984). Studies on the Epstein–Barr virus genome. Ph.D. thesis, University of Cambridge.

Glaser-Wuttke, G., Keppner, J. & Rasched, I. (1989). Pore-forming properties of the adsorption protein of filamentous phage fd. *Biochim. Biophys. Acta*, **985**, 239–247.

Güssow, D. & Clackson, T. (1989). Direct clone characterization from plaques and colonies by the polymerase chain reaction. *Nucl. Acids Res.* **17**, 4000.

Harlow, E. & Lane, D. (1988). *Antibodies–A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hermes, J. D., Parekh, S. M., Blacklow, S. C., Koster, H. & Knowles, J. R. (1989). A reliable method for random mutagenesis: the generation of mutant libraries using spiked oligodeoxyribonucleotide primers. *Gene*, **84**, 143–151.

Hoogenboom, H. R., Griffiths, A. D., Johnson, K. S., Chiswell, D. J., Hudson, P. & Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucl. Acids Res.* **19**, 4133–4137.

Hughes-Jones, N. C. & Gorick, D. B. (1991). Multiple epitopes recognized by human monoclonal IgM anti-D antibodies. *Vox Sanguinis*, **59**, 112–115.

Huse, W. D., Sastry, L., Iverson, S. A., Kang, A. S., Alting, M. M., Burton, D. R., Benkovic. S. J. & Lerner, R. A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science*, **246**, 1275–1281.

Huston, J. S., Levinson, D., Mudgett, H. M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Harer, E., Crea, R. & Oppermann, H. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 5879–5883.

Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. (1972). Vertebrate lysozymes. *The Enzymes*, Academic Press, New York and London.

Kang, A. S., Barbas, C. F., Janda, K. D., Benkovic, S. J. & Lerner, R. A. (1991). Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 4363–4366.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)*, **227**, 680–685.

Larrick, J. W., Danielsson, L., Brenner, C. A., Abrahamson, M., Fry, K. E. & Borrebaeck, C. A. (1989). Rapid cloning of rearranged immunoglobulin genes from human hybridoma cells using mixed primers and the polymerase chain reaction. *Biochem. Biophys. Res. Commun.* **160**, 1250–1256.

Liao, X. B. & Wise, J. A. (1990). A simple high-efficiency method for random mutagenesis of cloned genes using forced nucleotide misincorporation. *Gene*, **88**, 107–111.

Mäkelä, O., Kaartinene, M., Pelkonen, J. L. T. & Karjalainen, K. J. (1978). Inheritance of antibody specificity. Anti-2-Phenyloxazolone in the mouse. *J. Exp. Med.* **148**, 1644–1660.

Marks, J. D., Tristrem, M., Karpas, A. & Winter, G. (1991). Oligonucleotide primers for polymerase chain reaction amplification of human immunoglobulin variable genes and design of family-specific oligonucleotide probes. *Eur. J. Immunol.* **21**, 985–991.

McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature (London)*, **348**, 552–554.

Meyerhans, A., Vartanian, J. P. & Wain, H. S. (1990). DNA recombination during PCR. *Nucl. Acids Res.* **18**, 1687–1691.

Miller, J. H. (1972). *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Milstein, C. (1990). The Croonian lecture, 1989. Antibodies: a paradigm for the biology of molecular recognition. *Proc. Roy. Soc. Lond. Biol.* **239**, 1–16.

Mullinax, R. L., Gross, E. A., Amberg, J. R., Hay, B. N., Hogrefe, H. H., Kubitz, M., Greener, A., Alting, M. M., Ardourel, D., Short, J. M. & Shopes, R. (1990). Identification of human antibody fragment clones specific for tetanus toxoid in a bacteriophage lambda immunoexpression library. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 8095–8099.

Munro, S. & Pelham, H. R. B. (1986). An Hsp-like protein in the ER: identity with the 78 kD glucose regulated protein and immunoglobulin heavy chain binding protein. *Cell*, **46**, 291–300.

Orlandi, R., Güssow, D. H., Jones, P. T. & Winter, G. (1989). Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proc. Nat. Acad. Sci., U.S.A.* **86**, 3833–3837.

Parmley, S. F. & Smith, G. P. (1988). Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene*, **73**, 305–318.

Perelson, A. S. (1989). Immune network theory. *Immunol. Rev.* **110**, 5–36.

Persson, M. A. A., Caothien, R. H. & Burton, D. R. (1991). Generation of diverse high-affinity human monoclonal antibodies by repertoire cloning. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 2432–2436.

Roit, I. M., Brostoff, J. & Male, D. K. (1985). *Immunology*. Gower Medical Publishing Ltd, London.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1990). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Sanz, I., Kelly, P., Williams, C., Scholl, S., Tucker, P. & Capra, J. D. (1989). The smaller human VH gene families display remarkably little polymorphism. *EMBO J.* **8**, 3741–3748.

Schaaper, R. M. (1988). Mechanisms of mutagenesis in the *Escherichia coli* mutator mutD5: role of DNA mismatch repair. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 8126–8130.

Schibler, U., Marcu, K. B. & Perry, R. P. (1978). The synthesis and processing of the messenger RNAs specifying heavy and light chain immunoglobulins in MPC-11 cells. *Cell*, **15**, 1495–1509.

Scott, J. K. & Smith, G. P. (1990). Searching for peptide ligands with an epitope library. *Science*, **249**, 386–390.

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228**, 1315–1317.

Songsivilai, S., Bye, J. M., Marks, J. D. & Hughes, J. N. (1990). Cloning and sequencing of human lambda immunoglobulin genes by the polymerase chain reaction. *Eur. J. Immunol.* **20**, 2661–2666.

Thompson, K. M., Sutherland, J., Barden, G., Melamed, M. D., Randen, I., Natvig, J. B., Pascual, V., Capra, J. D. & Stebenson, F. K. (1991). Human monoclonal antibodies against blood group antigens preferentially express a VH4·21 variable region gene-associated epitope. *Scand. J. Immunol.* **34**, 509–518.

Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from the polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Nat. Acad. Sci., U.S.A.* **76**, 4350–4354.

Ward, E. S., Güssow, D., Griffiths, A. D., Jones, P. T. & Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature (London)*, **341**, 544–546.

Winter, G. & Milstein, C. (1991). Man-made antibodies. *Nature (London)*, **349**, 293–299.

Yamagishi, J., Kawashima, H., Matsuo, N., Ohue, M., Yamayoshi, M., Fukui, T., Kotani, H., Furuta, R., Nakano, K. & Yamada, M. (1990). Mutational analysis of structure–activity relationships in human tumor necrosis factor-alpha. *Protein Eng.* **3**, 713–719.

Edited by J. Karn

C18

MONOCLONAL ANTIBODIES TO HUMAN TUMOR NECROSIS FACTOR α: IN VITRO AND IN VIVO APPLICATION

Achim Möller,[1*] Franz Emling,[2] Dietmar Blohm,[3] Erich Schlick,[1] Klaus Schollmeier[1]

Three stable murine hybridoma cell lines, which secrete monoclonal antibodies (mAb) to human tumor necrosis factor α (TNFα), were established. None of the monoclonal antibodies cross-reacted with lymphotoxin, interleukin 2 (IL 2) or Interferon γ (IFNγ). The highly species-specific monoclonal antibody, designated as mAb 195, neutralizes the cytotoxic activity of human and chimpanzee TNFα. This antibody was further used during in vivo studies to neutralize human TNFα in a murine animal model. The mAb 114 is a weakly neutralizing antibody that binds to a different epitope of TNFα than mAb 195. mAb 114 shows a wide range of cross-reactivity with TNFα of the following species: dog, pig, cynomolgus, rhesus, baboon and chimpanzee. The mAb 199 binds to human TNFα, but does not neutralize the cytotoxic activity. The epitope recognized by this mAb is in close proximity to mAb 114. A reproducible, sensitive immunoassay for human TNF7α has been developed using the antibodies mAb 199 and mAb 195. The test is performed within 6 hr and detects TNF7α in serum samples, with a limit of detection of 5 to 10 pg/mL.

Tumor necrosis factor α (TNFα) was originally described as a serum factor produced in mice after an infection with Bacillus Calmette-Guérin and subsequently challenged with endotoxin.[1] It causes necrosis of Meth A sarcoma in mice. Subsequent studies showed the cytotoxic or cytostatic features of TNFα for a number of cell lines in vitro[2,3] and antitumoral effects in several in vivo murine tumor models.[4] Recent evidence shows that TNFα is a regulatory polypeptide and mediates inflammation and cellular immune responses.[5,6] It forms a complex network of interactive signals with other lymphokines as well as regulating production, growth, and differentiation of cells involved in inflammation, immunity, and hemopoiesis.[7]

Aggarwal, et al[8] described the purification of this cytotoxic factor from the human promyelocytic leukemia cell line HL60. Human TNFα is secreted as an unglycosylated 157-amino-acid polypeptide of relative molecular mass of 17350 Dalton. Lipopolysaccaride (LPS) is a potent trigger for TNFα production.[9,10,11] It is mainly secreted by macrophages but also produced by T lymphocytes and natural killer cells after stimulation.[12,13,14]

TNFα was cloned in *Escherichia coli* and a biologically active protein was expressed.[15] Using recombinant DNA techniques, we were able to produce large quantities of TNFα for crystallization experiments. Jones et al.[16] solved the structure of this lymphokine to a resolution of 2.9 Å.

TNFα has been linked to various toxic manifestations of infectious, neoplastic, or autoimmune diseases. For example, it has been implicated in the profound weight loss (cachexia) associated with chronic parasitic or bacterial diseases as well as with cancer.[5,17,18,19] The circulatory collapse and shock associated with acute bacterial infection,[20,21] the death of animals with cerebral malaria,[22] and the development of acute graft-versus-host disease and graft rejection[23,24] are other examples of the major involvement of TNFα in disease. Because TNFα plays a role in many life threatening diseases, the development of a molecule neutralizing the biological effects of TNFα would be of clinical value.

In this report, we describe the isolation of neutralizing and non-neutralizing monoclonal anti-TNFα antibodies. They bind to TNFα with higher affinity than any of

[1]BASF Bioresearch Corp., Cambridge, MA,
[2]Knoll AG, Department of Oncology, Ludwigshafen, FRG
[3]BASF AG, Hauptlaboratorium, Ludwigshafen, FRG
*To whom reprint requests should be addressed.

1043-4666/90/0203-0002$05.00/0

KEY WORDS: Monoclonal antibodies/Tumor necrosis factor α/ELISA/In vivo use

the previously described monoclonal anti-TNFα antibodies. Furthermore, the mAb 195 shows a highly restricted species specificity not described before. This antibody has been used at equimolar concentrations to neutralize the biological effects of human TNFα under in vitro and in vivo conditions. A combination of the antibodies mAb 199 and mAb 195 was used to develop a highly sensitive sandwich immunoassay.

RESULTS

Production of Monoclonal Antibodies

Spleen cells from mice showing high titers of antibody in their sera were fused with SP2/0-Ag14 cells. Two weeks after fusion, culture supernatants were harvested from each well and assayed for the secretion of anti-TNFα antibodies using an enzyme immunoassay and the inhibition of cytolytic activity as described in the Materials and Methods section. Twelve positive cultures were identified and after two subcloning steps performed by limiting dilution, three monoclonal anti-TNFα antibodies were further investigated. We designated these monoclonal antibodies mAb 114, mAb 195 and mAb 199. The general characteristics of these monoclonal antibodies are shown in Table 1.

Epitope Mapping of TNFα Using Monoclonal Antibodies

The number of distinct TNFα antigenic epitopes, defined by the panel of monoclonal antibodies, was determined by an enzyme immunoassay that used an antibody competition method. Three epitopes were identified, two of which reacted with the neutralizing antibodies mAb 114 and mAb 195.

Fig. 1 depicts the inhibition experiments showing the competition between biotinylated mAb 114(a), mAb 195(b) or mAb 197(c) and the unlabeled mAbs. In all cases tested, the binding of the labeled mAb was completely inhibited by the corresponding unlabeled mAb. The data shown in Fig. 1 suggest that the epitopes recognized by mAb 114 and mAb 199 are in close proximity. mAb 195 did not compete with the other mAbs tested, suggesting that the epitope recognized by mAb 195 is located at a further distance than the other mAbs. mAb 114 has an influence on the mAb 195 binding site but not vice versa. As TNFα can be detected on Western blots by the three different mAbs, this leads

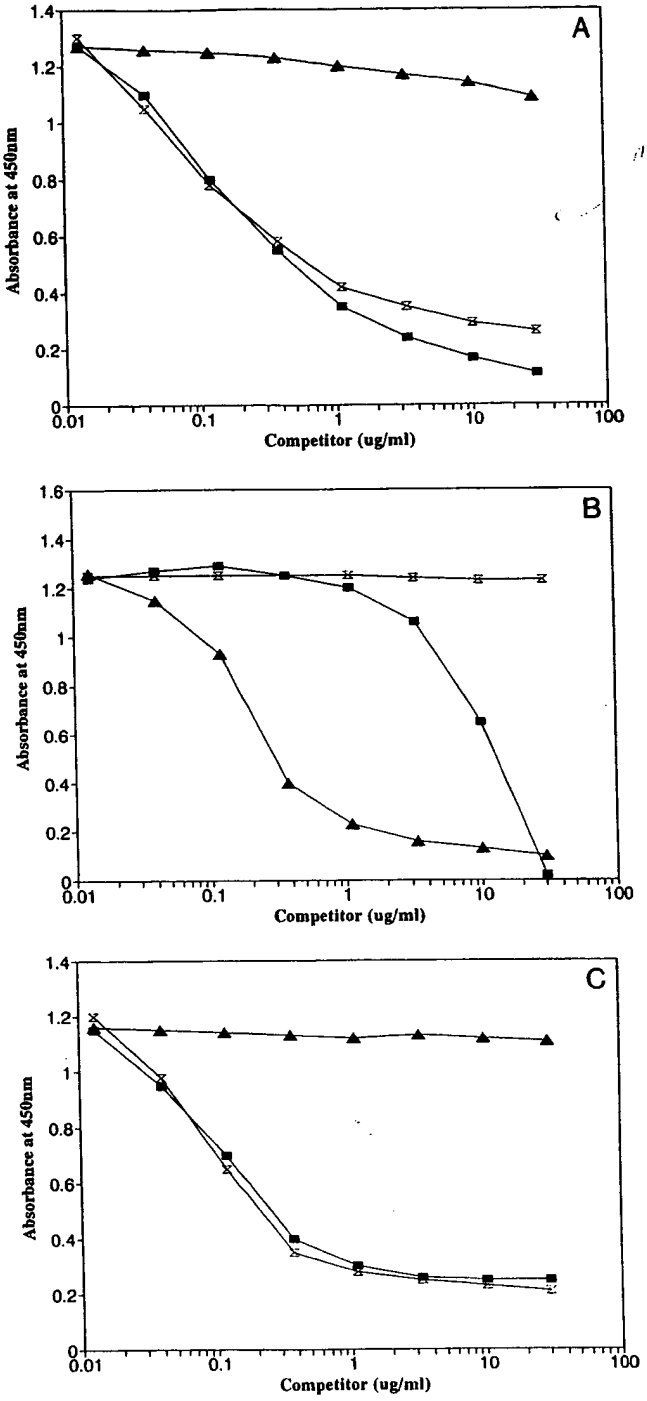


Figure 1. Epitope mapping of TNFα.

Binding of biotin-labeled monoclonal antibodies (A-C) to recombinant human TNFα was competed by unlabeled monoclonal antibodies, using as competitors unlabeled mAb 114 (closed square, ■), mAb 195 (closed triangle, ▲) and mAb 199 (open hourglass, ⧖).

TABLE 1. General characteristics of the anti-human TNFα specific monoclonal antibodies mAb 114, mAb 195 and mAb 199

	mAb 114	mAb 195	mAb 199
Isotype	G1, κ	G3, κ	G1, κ
Affinity Constant x 10^9 L/mole	1.4	3.5	2.0

The isotypes of the various mAbs were determined by specific reagents from Cappel. The determination of the affinity constants was performed with a non-radioactive immunoassay using nonlinear-least-square-fit models.

CYTOKINE, Vol. 2, No. 3 (May 1990: 162-169)

to the speculation that all three epitopes consist of a sequential stretch of amino acids on the TNFα molecule.

Characterization of Monoclonal Antibodies

The mAbs were tested for binding to TNFα and related cytokines by enzyme immunoassay, and for neutralization of cytolytic activity. Both natural and recombinant TNFα were recognized equally well, while the monoclonal antibodies did not crossreact with Lymphotoxin (TNFβ), IL2, or IFNγ (data not shown).

The studies for neutralization of TNFα cytolytic activity were done in the L929 assay and are shown in Fig. 2. The best neutralizing antibody, mAb 195, had a neutralization titer of 1,000 U TNFα neutralized/μg of purified IgG. Under the assumption that the TNFα molecule forms a trimeric complex in solution,[25] the activity of one TNFα molecule can be completely neutralized by three mAbs. The TNFα neutralizing activity of mAb 114 is reduced by a factor of 20 when compared to the activity of mAb 195. mAb 199, on the other hand, is able to neutralize the activity of TNFα only at extremely high concentrations (neutralization titer of 1.6 U TNFα neutralized/μg of purified IgG), defining mAb 199 as non-neutralizing. Human Lymphotoxin (TNFβ) shows the same cytotoxic effects on L929 cells in vitro as TNFα. Neither of the TNFα specific monoclonal antibodies mAb 114 and mAb 195 neutralized TNFβ, even when tested at concentrations greater than two hundred times the amount required to neutralize TNFα.

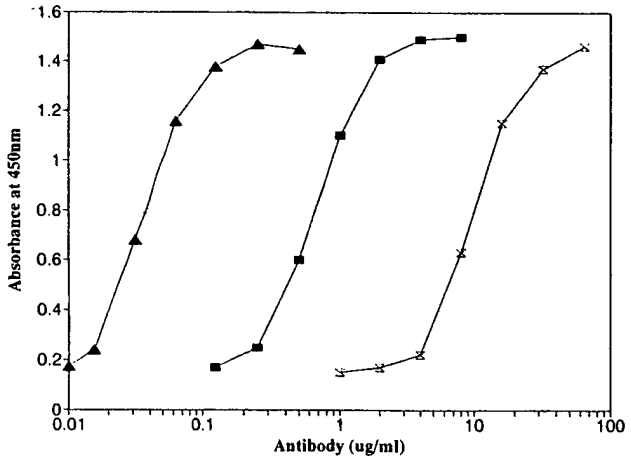


Figure 2. Neutralization of the cytotoxic activity of human TNFα by mAbs 114, 195 and 199.

A confluent monolayer of L929 cells is completely lysed by 12.5 ng/mL TNFα after 48 hr. In the neutralization experiments, the monclonal antibodies mAb 114 (closed square, ■), mAb 195 (closed triangle, ▲) and mAb 199 (open hourglass, ⧖) were incubated with antigen for 2 hr at room temperature prior to the assay. This time was sufficient to completely neutralize TNFα cytolytic activity.

TABLE 2. Neutralization of the cytotoxic activity of TNFα from different animal species

	mAb 114	mAb 195
Mouse	–	–
Rat	–	–
Rabbit	–	–
Dog	+	–
Pig	+	–
Cynomolgus	+	–
Rhesus	+	–
Baboon	+	–
Chimpanzee	+	+
Human	+	+

The TNFα activities in the supernatants of peripheral mononuclear cells of the different animal species were determined by the L929 cytotoxic assay. For the neutralization studies a dilution was chosen, which gave more than 80% cytotoxicity. Two-fold serial dilutions of monoclonal antibodies were prepared and incubated with the different supernatants for 2 hr at room temperature before the assay.

Cross-Reactivity of Monoclonal Antibodies with TNFa from Different Animal Species

The cross-reactivities of the monoclonal antibodies were tested by their abilities to neutralize TNFα in different species (Table 2). mAb 195 is highly specific. This is shown by its neutralization of only human and chimpanzee TNFα. mAb 114 neutralizes TNFα from other animal species, but does not react with mouse, rat, and rabbit TNFα.

Characterization of Monoclonal Antibodies by Immunoblot

The monoclonal antibodies were examined for their binding to sodium dodecyl sulfate (SDS) denatured TNFα in the presence of human serum. This was done using the immunoblot technique.[26] A representative blot is shown for mAb 195 in Fig. 3. The antibody bound the monomeric form of TNFα and did not show any cross-reactivity with human proteins. The same results were found for mAbs 114 and 199 (data not shown). These results indicate that the three monoclonal antibodies probably bind to sequential epitopes, which are not disrupted by treatment with SDS.

Quantification of TNFa Using an ELISA Method

A specific sandwich ELISA was developed as a potential diagnostic assay in order to monitor serum levels of TNFα in patients. Purified mAb 199 was adsorbed onto plastic micro-ELISA wells and used to specifically bind recombinant or natural TNFα. Biotinylated mAb 195, followed by the streptavidin-peroxidase complex, was then used to detect bound antigen. After incubation with TMB solution, the absorbance of each well was measured with an automated SLT Easy reader. A typical standard curve is shown in Fig. 4. The same curve can be seen when TNFα is titrated in 25% human

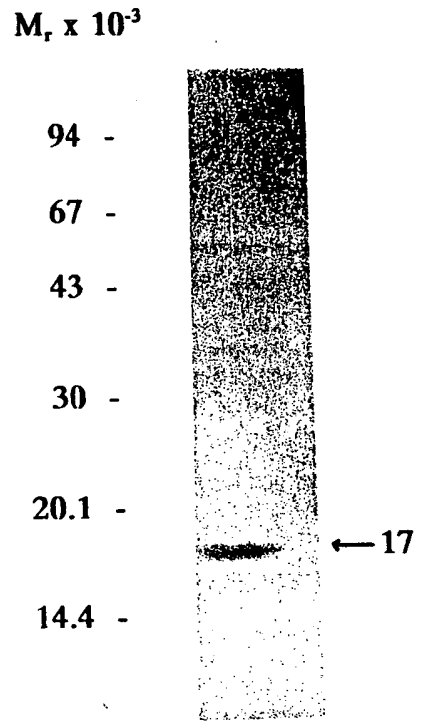


Figure 3. TNFα immunoblot probed with mAb 195.

Recombinant TNFα (0.2 μg) was mixed with human serum and chromatographed on a 15% Laemmli gel. The gel was blotted and tested with mAb 195. The position of the TNFα is shown by the arrow, designated A, which indicates the estimated molecular weight in Daltons x 10^3.

serum (data not shown). The assay has a limit of detection of 5 to 10 pg/mL (2 times the standard deviation) and a linear range up to 300 pg/mL. The specificity was demonstrated by the lack of any signal generated by 10 μg/mL TNFβ, IFNγ and IL2.

In order to compare the results obtained from ELISA with the bioassay, we measured TNFα levels in column fractions from a typical purification, using both types of assays. The ratio of bioactive to immunoreactive TNFs (i.e., the specific activity) of all of the TNFα fractions was relatively constant (data not shown). This result demonstrates that there was no significant difference in the levels of TNFα that were measured by ELISA or by bioassay.

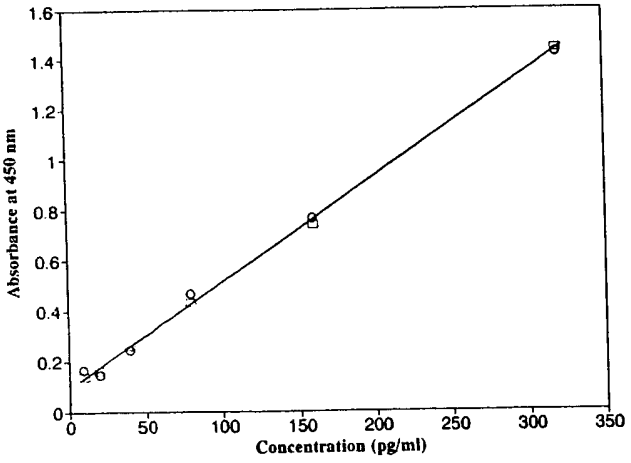


Figure 4. Standard curve for recombinant and natural TNFα ELISA.

The open squares represent recombinant and the open circles natural TNFα. Up to 10 μg/mL TNFβ, IFNγ and IL 2 gave no signal.

Neutralization of Human TNFα in the Mouse

To investigate the potential in vivo use of the monoclonal antibodies, Balb/c mice were treated with a lethal dose of TNFα. The toxicity could be completely neutralized by mAb 195, as shown in Fig. 5. Tested at a 0.3 M ratio to TNFα, the mAb 195 demonstrated only a minor effect on the survival of the mice; while given at a 1:1.5 M ratio, mAb 195 completely blocked the toxic effects of human TNFα in the mouse. The control antibody, mAb 199, had no effect on human TNFα activity in the mice and neither monoclonal antibody mAb 195 nor monoclonal antibody mAb 199 alone had a toxic effect on the animals.

DISCUSSION

In the present study, we have described the generation of three high affinity murine monoclonal antibodies to human TNFα, which we have designated mAb 114, 195, and 199. Although monoclonal antibodies against TNFα have previously been described,[20,27,28,29,30,31,32] the mAb 199 and, in particular, mAb 195 that are obtained here, are quite different and useful for evaluating the functions of TNFα under in vitro and in vivo conditions.

We isolated two types of neutralizing monoclonal anti-TNFα antibodies that bind to distinct antigenic epitopes. The epitope recognized by mAb 195 is re-

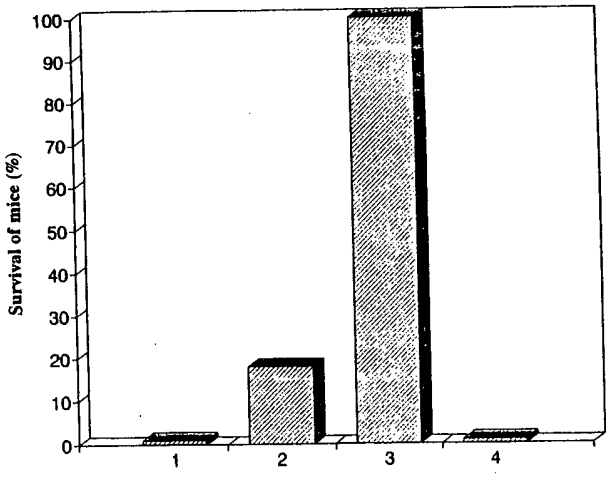


Figure 5. Neutralization of human TNFα in the mouse.

Survival of Balb/c mice was determined (1) 24 hr after the injection of TNFα (2); TNFα + mAb 195, ratio 1:0.3 (3); TNFα + mAb 195, ratio 1:1.5, and (4) TNFα + mAb 199.

stricted to only human and chimpanzee TNFα. This antibody equally recognizes recombinant and native TNFα and does not show any cross-reactivity with such lymphokines as lymphotoxin, IL 2, and IFNγ. The high binding constant and the above-mentioned specificity enabled us to use this antibody efficiently in neutralization studies. Under in vivo and in vitro conditions, about 3 molecules of antibody per molecule of TNFα are sufficient for blocking the lethal or cytotoxic effects. Furthermore, in our murine system, we could apply the monoclonal antibody, mAb 195, even after the injection of TNFα and still find complete survival of the mice. Fendly et al.[31] postulated that mAbs neutralize the cytotoxicity of TNFα by blocking the binding of TNFα to its receptor. By fluorescent activated cell sorter (FACS) analysis it was shown that the monoclonal antibody mAb 195 can still bind when TNFα occupies the receptor. This suggests an additional mechanism for neutralization of the biological activity of TNFα (Heilig B, Möller A, and Dörken B, manuscript submitted for publication). The TNFα molecule can then be divided into a receptor binding and a neutralization area. The second neutralizing antibody, mAb 114, can bind simultaneously with mAb 195 to TNFα. The epitope recognized by this monoclonal antibody is related spatially to mAb 199, the non-neutralizing antibody. These results support the above model that the biological activity of TNFα is not restricted to one unique site.

The monoclonal antibodies mAb 195 and mAb 199 were used to develop a sensitive sandwich ELISA assay. The ELISA was as sensitive as the bioassay and could discriminate between TNFα and TNFβ, which the bioassay is unable to do. Furthermore, it is faster and easier to perform than the bioassay.

The monoclonal antibodies, described here, should prove to be useful in the study of this cytokine's mechanisms of actions. The sensitive ELISA assay can be used to identify and quantitate TNFα in human diseases and for the monitoring of TNFα levels in patients undergoing TNFα treatment. The neutralizing antibodies can selectively neutralize TNFα cytolytic activity in an in vitro system, thereby separating the effects of TNFα from other cytokines in complex biological systems. Neutralizing anti-TNFα monoclonal antibodies may be able to block the side effects of TNFα in vivo, as shown by Tracey et al. 1987,[33] for septic shock during lethal bacteremia in baboons.

Studies with polyclonal anti-TNFα antibodies in mice have also shown that this treatment almost completely prevented the cutaneous and intestinal lesions of the acute-phase of Graft-versus-Host Disease and markedly reduced overall mortality.[23] In the studies of Grau et al.[22] the injection of a rabbit antibody to TNFα fully protects mice from cerebral malaria. TNFα-specific monoclonal antibodies may therefore provide a specific therapy for TNFα-mediated disorders.

MATERIALS AND METHODS

Recombinant Lymphokines

Recombinant human and murine TNFα were produced in *E. coli* and purified to 8×10^6 and 13×10^6 laboratory U/mg, respectively.[34] The materials appeared homogeneous when analyzed by SDS-PAGE, stained with Coomassie Blue. Natural human TNFα was purified from HL-60 cell line supernatants.[8] Recombinant human IFNγ and human lymphotoxin (TNFβ) were produced in *E. coli* and purified to homogeneity.

TNFα from Different Animal Species

Blood was drawn from the following animal species: mouse, rat, rabbit, dog, pig, cynomolgus, rhesus, baboon and chimpanzee, and centrifuged through a ficoll gradient. Peripheral mononuclear cells (2×10^6 cells/mL) were suspended in RPMI 1640 supplemented with penicillin (50 U/mL), streptomycin (50 μg/mL), glutamine (2 mM) and 5% fetal calf serum (FCS) (Boehringer). 10 mL of this suspension was transferred to a T25 Falcon flask. TNFα was then induced by adding 10 μg/mL LPS (Sigma, No. L4005). The supernatants were collected after 16 hr and the TNFα levels were either immediately estimated or stored at −80°C until the TNFα test was performed.

Biological Assay for TNFα Activity

The biological activity of TNFα was determined by the murine L929 fibroblast cytotoxic assay, as described by Aggarwal et al.[8] Briefly, L929 cells were seeded at a density of 10,000 cells per well in 96-well microtiter plates (Nunc) using 0.1 mL Eagle's Minimal Essential Medium that contained 10% fetal bovine serum (FBS). An equal volume of medium containing the desired dilutions of TNFα, with or without anti-TNFα antibodies, was added and incubated at 37°C and 5% CO_2 for 48 hr in a humidified incubator. The cells were then stained with 0.5% crystal violet. After an incubation of 15 min, the plates were rinsed with tap water and then dried. Crystal violet was dissolved by the addition of 0.1 mL/well of 0.1% acetic acid in 50% ethanol. The absorbance at 580 nm was read with a SLT Easy Reader. The TNFα U/mL represent the reciprocal of the TNFα dilution that causes 50% cytotoxicity under the conditions of the assay.

Immunization, Fusion and Cloning

Purified recombinant human TNFα, in an emulsion formed by equal volumes of TNFα solution and Freund's complete adjuvant, was injected into BALB/c mice i.p. (5 μg/mouse). Two, 4 and 6 weeks later, the mice received the same amount of antigen in Freund's incomplete adjuvant. Three days after the last immunization spleen cells from mice showing high titers of antibody in their sera were fused with SP2/0-Ag14 cells at a ratio of 5:1, according to previously described techniques.[35] Two weeks after fusion, culture supernatants were assayed for secretion of anti-TNFα antibodies using the methods described. Positive hybridomas were subcloned twice by limiting dilution in 96-well plates, using BALB/c thymocytes (10^6 cells/well) as feeder cells.

Selection of Monoclonal Antibodies to Human TNFα

Recombinant human TNFα was diluted in PBS to a concentration of 0.01 mg/mL. A 0.1 mL aliquot of this suspension was added to each well of a 96-well microtiter plate (Flow). The plate was incubated overnight at 4°C and the fluid was removed. Residual protein-binding sites were blocked by PBS containing 0.1% Tween 20 (PBS-T). Supernatants (0.1 mL/well), which contained monoclonal antibodies, were added and incubated for 2 hr at room temperature. After a washing step, 0.1 mL of a 1,000-fold diluted solution of goat anti-mouse IgG antibody labeled with horseradish peroxidase (Boehringer) was added to each well. After incubation for 1 hr at room temperature and washing with PBS-T, 0.1 mL of 0.42 mM 3,3',5,5'-Tetramethylbenzidine (TMB, Boehringer Mannheim) in 0.1 M acetate-citrate buffer, pH 4.9 containing 0.004% H_2O_2 was added. The reaction was stopped after 10 min by the addition of 2.0 N sulfuric acid and the absorbance at 450 nm of each well was determined using a SLT Easy Reader.

Monoclonal Antibody Isotyping

A 96-well microtiter plate (Flow) was coated with human rTNFα (0.01 mg/mL in PBS) overnight at 4°C and then blocked with PBS containing 1% bovine serum albumin (BSA). 0.1 mL of culture supernatants were added to wells and incubated for 1 hr at room temperature. Plates were washed with PBS-T and 0.05 mL of each isotype-specific rabbit anti-mouse Ig antibody (Cappel) was added. After 1 hr incubation at room temperature, the wells were washed 3 times and 0.05 mL of horseradish peroxidase labeled goat anti-rabbit IgG (Boehringer Mannheim) was added to each well. Incubation was continued for another hour at room temperature. Plates were washed 5 times with PBS-T and 3, 3', 5, 5'-Tetramethylbenzidine (TMB) was used as substrate as described previously.

Determination of Epitope Specificities

A 96-well microtiter plate (Flow) was coated with 0.001 mg/mL recombinant human TNFα (0.1 mL/well) in PBS overnight at 4°C. Nonspecific binding sites were then blocked with a solution of 1% BSA in PBS for 0.5 hr at 37°C. A constant volume (0.05 mL) of an appropriate dilution of biotinylated monoclonal antibody was added to equal volumes of double dilutions of respective unlabeled monoclonal antibodies. The mixture was then added to the wells and incubated for 2 hr at 37°C. The washing and further processing were done as described for "ELISA."

Biotinylation of Monoclonal Antibodies

The mAbs were biotinylated with D-biotinyle-e-amidocaproic-acid-N-hydroxysuccinimide-ester (Boehringer Mannheim). Ten mg of this reagent was dissolved in 1 mL DMSO. The monoclonal antibodies were adjusted to 5 mg/mL in PBS. A 35 molar excess of biotin was used to label the antibodies for 2 hr at room temperature. The unbound biotin was removed from modified monoclonal antibodies by Sepharose G25 column chromatography (1 × 15 cm). The column was run with 0.05 M potassium phosphate, 0.5 M sodium chloride and 5% saccharose at pH 7.8.

Purification of Monoclonal Antibodies

Hybridoma cells were cultivated in Iscove Medium containing the following supplements: 1.2 mg/mL BSA, 0.024 mg/mL soybean lipids, 0.034 mg/mL transferrin and 5×10^{-5} M β-mercaptoethanol. Cells were grown in spinner flasks, the medium was harvested by filtration through a 0.3 μm filter (Pall) and concentrated 10-fold by ultrafiltration using a PM10 membrane (Amicon). The concentrate was dialyzed against 1.5 M glycine and 3 M NaCl, pH 8.9 and the IgG purified by Protein A Sepharose chromatography (Pharmacia), according to the manufacturer's recommendation.

Monoclonal Antibodies Affinity Determination

96-well microtiter plates (Nunc) were coated for 16 hr at 4°C with 0.01 mg/mL TNFα (0.1 mL/well) in PBS. After that the wells were washed with PBS-T and post-coating was done with 1% BSA in PBS for 0.5 hr at 37°C. mAbs were serially diluted with 0.1% BSA in PBS and incubation was done for 4 hr at 37°C. The wells were washed again 5 times with PBS containing 0.1% Tween 20 (PBS-T) and incubated for 2 hr at 37°C with horseradish peroxidase conjugated rabbit anti-mouse IgG (0.1 mL/well, diluted 1:3,000 in 0.1% BSA in PBS). The wells were then washed 5 times with PBS-T and TMB was used as substrate, as described previously.

Calculations of the affinity of the monoclonal antibodies were done by non-linear least square fits. These were based on the model of one monoclonal antibody binding site per TNFα subunit.

Enzyme-Linked Immunosorbent Assay (ELISA)

A 96-well microtiter plate (Nunc) was coated overnight with 0.1 mL/well of anti-TNFα mAb 199 (0.005 mg/mL) in PBS. Post coating was done with 1% BSA in PBS for 0.5 hr at 37°C. The plate was washed 5 times with PBS containing 0.1% Tween 20 (PBS-T). For a standard curve, TNFα was diluted in 20 mM Tris-HCl, 150 mM NaCl and 1% BSA, pH 7.4. TNFα-containing medium or serum (0.1 mL/well) was reacted for 3 hr at room temperature or overnight at 4°C. After washing, 0.1 mL/well of biotinylated mAb 195 (0.001 mg/mL) was added and incubated for 2 hr at room temperature. After 3 washes with PBS-T, 0.1 mL/well of 10,000-fold diluted streptavidin-peroxidase complex (Boehringer Mannheim) was added and incubated again for 0.5 hr at room temperature. TMB was used as a substrate for developing the color as described.

Gel Electrophoresis and Immunoblots

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed, as described by Laemmli et al.[36] Recombinant TNFα (0.2 μg) and human serum were subjected to electrophoresis on a 15% acrylamide gel. Gels were stained with 0.15% Coomassie Blue R250 (wt/vol) in 25% isopropanol (vol/vol), 10% acetic acid (vol/vol), and destained with 10% methanol (vol/vol), 10% acetic acid (vol/vol). The binding of antibodies to TNFα, which are

separated by SDS-PAGE, was assessed by the immunoblot technique.

SDS-PAGE gels were transferred to nitrocellulose as previously described by Towbin et al.[26] The non-specific reactive sites on the nitrocellulose were blocked by incubating for 30 min with 3% gelatin (wt/vol) in TBS. Test monoclonal antibodies were diluted to 0.01 mg/mL in 1% gelatin (wt/vol) in TBS, and incubated with the nitrocellulose strips for 16 hr at 22°C. The strips were then washed for 15 min, each with three changes of PBS-T. Bound monoclonal antibodies were detected after incubating for 4 hr with horseradish peroxidase conjugated anti-murine IgG (Cappel) that was diluted 1:1000 in 1% gelatin. The strips were washed as above and incubated 15 min with 4-chloro-1-naphthol (0.6 mg/mL in 20% methanol, 80% TBS containing 0.018% H_2O_2), rinsed with water and air dried.

In Vivo Neutralization of Human TNFα

Four-6 week old BALB/c mice were randomized and divided into groups of 5. The substances were given i.v. into the lateral tail vein; the volume of injection did not exceed 10 mL/kg. TNFα (2 mg/kg body weight) was applied first, followed 15 to 30 min later by the monoclonal antibodies (doses: 2 mg/kg and 10 mg/kg). The rate of survival was determined 24 hr later.

Protein Determination

Protein determinations were performed by the dye binding method of Bradford,[37] which uses bovine IgG as a standard. The protein concentration was also determined by absorbance using the extinction coefficients for pure recombinant TNFα, ϵ_{280} 1 mg/mL = 1.6.

Acknowledgments

The authors thank M. Pan for help in writing the manuscript, J. Delzer and H. Hillen for fermentation and purification of the cytokines. We also gratefully acknowledge the expert technical assistance of M. Seitz, H. Hofmann, E. Klimm and R. Merx.

REFERENCES

1. Carswell EA, Old LJ, Kassel RL, Greene S, Fiore N, Williamson B (1975) An endotoxin-induced serum factor that causes necrosis of tumors. Proc Nat Acad Sci USA 72:3666-3670

2. Haranaka K, Satomi N (1981) Cytotoxic activity of tumor necrosis factor on human cancer cells in vitro. Jpn J Exp Med 51:191-194

3. Sugarman BJ, Aggarwal BB, Hass PE, Figari IS, Palladino MA, Jr, Shepard HM (1985) Recombinant human tumor necrosis factor-α: Effects of proliferation of normal and transformed cells in vitro. Science 230:943-949

4. Haranaka K, Satomi N, Sakurai A (1984) Antitumor activity of murine tumor necrosis factor against transplanted murine tumors and heterotransplanted human tumors in nude mice. Int J Cancer 34:263-267

5. Beutler B, Cerami A (1987) Cachectin: More than a tumor necrosis factor. N Engl J Med 316:379-383

6. Kunkel SL, Remick DG, Strieter RM, Larrick JW (1989) Mechanisms that regulate the production and effects of tumor necrosis factor-α. Critical Rev Immunol 9:93-117

7. Old LJ (1987) Polypeptide mediator network. Nature 326:330-331

8. Aggarwal BB, Kohr WJ, Hass PE, Moffat B, Spencer SA, Henzel WJ, Bringman TS, Nedwin GE, Goeddel DV, Harkins RN (1985) Human tumor necrosis factor, production, purification characterization. J Biol Chem 260:2345-2354

9. Beutler B, Milsark IW, Cerami AC (1985) Passive immunization against cachectin/tumor necrosis factor protects mice from lethal effects of endotoxin. Science 229:869-871

10. Chen AR, McKinnon KP, Koren HS (1985) Lipopolysaccharide (LPS) stimulates fresh human monocytes to lyse actinomycin D-treated WEHI-164 target cells via increased secretion of a monokine similar to tumor necrosis factor. J Immunol 135:3978-3984

11. Michie HR, Manogue KR, Spriggs DR, Revhaug A, O'Dwyer S, Dinarello CA, Cerami A, Wolff SM, Wilmore DW (1988) Detection of circulating tumor necrosis factor after endotoxin administration. N Engl J Med 318:1481-1486

12. Cuturi MC, Murphy M, Costa-Giomi MP, Weinmann R, Perussia B, Trinchieri G (1987) Independent regulation of tumor necrosis factor and lymphotoxin production by human peripheral blood lymphocytes. J Exp Med 165:1581-1585

13. Feinman R, Henriksen-DeStefano D, Tsujimoto M, Vilcek J (1987) Tumor necrosis factor is an important mediator of tumor cell killing by human monocytes. J Immunol 138:635-641

14. Debets JMH, Van der Linden CJ, Spronken IEM, Buurman WA (1988) T-cell mediated production of tumor necrosis factor α by monocytes. Scand J Immunol 27:601-608

15. Pennica D, Nedwin GE, Hayflick JS, Seeburg PH, Derynck R, Palladino MA, Kohr WJ, Aggarwal BB, Goeddel DV (1984) Human tumour necrosis factor: Precursor structure, expression and homology to lymphotoxin. Nature 312:724-729

16. Jones EY, Stuart DI, Walker NPC (1989) Structure of tumor necrosis factor. Nature 338:225-228

17. Cerami A, Ikeda Y, Le Trang N, Hotez PJ, Beutler B (1985) Weight loss associated with an endotoxin-induced mediator from peritoneal macrophages: The role of cachectin (tumor necrosis factor) Immunol Lett 11:173-177

18. Scuderi P, Sterling KE, Lam KS, Finley PR, Ryan KJ, Ray CG, Petersen E, Slymen DJ (1986) Raised serum levels of tumour necrosis factor in parasitic infections. Lancet 2:1364-1365

19. Shimomura K, Manda T, Mukumato S, Kobayashi K, Nakano K, Mori J (1988) Recombinant human tumor necrosis factor α: Thrombus formation as a cause of anti-tumor activity. Int J Cancer 41:243-247

20. Tracey KJ, Beutler B, Lowry SF, Merryweather J, Wolpe S, Milsark IW, Hariri RJ, Fahey TJ III, Zentella A, Albert JD, Shires GT, Cerami A (1986) Shock and tissue injury induced by recombinant human cachectin. Science 234:470-474

21. Waage A, Halstensen A, Espevik T (1987) Association between tumour necrosis factor in serum and fatal outcome in patients with meningococcal disease. Lancet 1:355-357

22. Grau GE, Fajardo LF, Piguet P-F, Allet B, Lambert P-H, Vassalli P (1987) Tumor necrosis factor (Cachectin) as an essential mediator in murine cerebral malaria. Science 237:1210-1212

23. Piguet P-F, Grau GE, Allet B, Vassalli P (1987) Tumor necrosis factor/cachectin is an effector of skin and gut lesions of the acute phase of graft-vs-host disease. J Exp Med 166:1280-1289

24. Maury CPJ, Teppo AM (1987) Raised serum levels of cachectin/tumor necrosis factor a in renal allograft rejection. J Exp Med 166:1132-1137

25. Smith RA, Baglioni C (1987) The active form of tumor necrosis factor is a trimer. J Biol Chem 262:6951-6955

26. Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc Natl Acad Sci USA 76:4350-4354

27. Bringman T, Aggarwal BB (1987) Monoclonal antibodies to human tumor necrosis factors alpha and beta: Application for affinity

purification, immunoassays, and as structural probes. Hybridoma 6:489-507

28. Hirai M, Okamura N, Terano Y, Tsujimoto M, Nakazato H (1987) Production and characterization of monoclonal antibodies to human tumor necrosis factor. J Immunol Methods 96:57-62

29. Meager A, Parti S, Leung H, Peil E, Mahon B (1987) Preparation and characterization of monoclonal antibodies directed against antigenic determinants of recombinant human tumour necrosis factor (rTNF). Hybridoma 6:305-311

30. Liang CM, Liang SM, Jost T, Sand A, Dougas I, Allet B (1986) Production and characterization of monoclonal antibodies against recombinant human tumor necrosis factor/cachectin. Biochem Biophys Res Comm 137:847-854

31. Fendly BM, Toy KJ, Creasey AA, Vitt CR, Larrick JW, Yamamoto R, Lin LS (1987) Murine monoclonal antibodies defining neutralizing epitopes on tumor necrosis factor. Hybridoma 6:359-369

32. Shimamoto Y, Chen RL, Bollon A, Chang A, Khan A (1988) Monoclonal antibodies against human recombinant tumor necrosis factor: Prevention of endotoxic shock. Immunol Lett 17:311-318

33. Tracey KJ, Fong Y, Hesse DG, Manogue KR, Lee AT, Kuo GC, Lowry SF, Cerami A (1987) Anti-cachectin/TNF monoclonal antibodies prevent septic shock during lethal bacteraemia. Nature 330:662-665

34. Ziegler-Heitbrock HW, Möller A, Linke RP, Haas JG, Rieber EP and Riethmüller (1986) Tumor necrosis factor as effector molecule in monocyte mediated cytotoxicity. Cancer Res 46:5947-5952

35. Oi V, Herzenberg L (1980) Immunoglobulin-producing hybrid cell lines. In Mishell B, Shiigi S (eds). Selected Methods in Cellular Immunology. Freeman Co., San Francisco, pp 351-372.

36. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685

37. Bradford, MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-254

The EMBO Journal vol.14 no.18 pp.4628–4638, 1995

The structural repertoire of the human V_κ domain

Ian M.Tomlinson[1], Jonathan P.L.Cox[2], Ermanno Gherardi[3], Arthur M.Lesk[4] and Cyrus Chothia[1,5]

[1]MRC Centre for Protein Engineering, [5]MRC Laboratory of Molecular Biology, [3]ICRF Cell Interactions Laboratory, [4]Department of Haematology, University of Cambridge, Hills Road, Cambridge CB2 2QH and [2]School of Chemistry, University of Bath, Claverton Down, Bath, Avon BA2 7AY, UK

In humans, the gene for the V_κ domain is produced by the recombination of one of 40 functional V_κ segments and one of five functional J_κ segments. We have analysed the sequences of these germline segments and of 736 rearranged V_κ genes to determine the repertoire of main chain conformations, or canonical structures, they encode. Over 96% of the sequences correspond to one of four canonical structures for the first antigen binding loop (L1) and one canonical structure for the second antigen binding loop (L2). Junctional diversity produces some variation in the length of the third antigen binding loop (L3) and in the identity of residues at the V_κ–J_κ join. However, this is limited and 70% of the rearranged sequences correspond to one of three known canonical structures for the L3 region. Furthermore, we show that the canonical structures selected during the primary response are conserved during affinity maturation: the key residues that determine the conformations of the antigen binding loops are unmutated or undergo conservative mutation. The implications of these results for immune recognition are discussed.

Keywords: antibodies/canonical structures/human/kappa/V_κ

Introduction

The antigen binding site of an antibody is formed by six loops of polypeptide: three from the light chain variable domain (V_L) and three from the heavy chain variable domain (V_H) (see Figure 1). A diverse repertoire of V genes that encode the V_L and V_H domains is produced by the combinatorial rearrangement of gene segments (Tonegawa, 1983), that are drawn from pools of moderate size. The V_L gene is produced by the recombination of two gene segments, V_L and J_L. The V_L segment codes for residues 1–95 which include the regions that form the first and second antigen binding loops. The third antigen binding loop is formed by the end of the V_L segment and the beginning of the J_L segment. Antibodies encoded by these rearranged genes are believed to be sufficiently diverse to recognize almost all antigens with at least moderate affinity. High-affinity antibodies are produced by somatic mutation of the rearranged genes, followed by selection for improved binding (Berek and Milstein, 1987).

V_L domains are divided into two classes: kappa (κ) and lambda (λ). Work over the past 15 years has resulted in what is probably a complete description of the sequences that encode the human V_κ domain [see Zachau (1993) for a review]. The V_κ segments are arranged in two separate regions on chromosome 2, named after their positions relative to the five J_κ segments. The J_κ-distal region, that contains 36 V_κ segments, is a partial duplicate of the larger J_κ-proximal region, that contains 40 V_κ segments.

Analysis of the structures and sequences of antibodies from humans and mice has shown that the three antigen binding loops of the V_κ domain have a small repertoire of main chain conformations, or canonical structures, that is largely shared by both species (Chothia and Lesk, 1987; Chothia *et al.*, 1989; Brünger *et al.*, 1991; He *et al.*, 1992). The canonical structures are determined by (i) the length of the antigen binding loop and (ii) particular residues, or types of residues, at key sites in the loop itself and in the antibody framework. These sequence features have been described in detail (see the references above) and can be used to predict the canonical structures of the three antigen binding loops of the V_κ domain.

In this paper, we describe the canonical structures implicit in human V_κ sequences. To describe the structural repertoire available to V_κ domains in the primary repertoire, the sequences of the germline V_κ segments are first examined. We go on to look at 736 rearranged V_κ

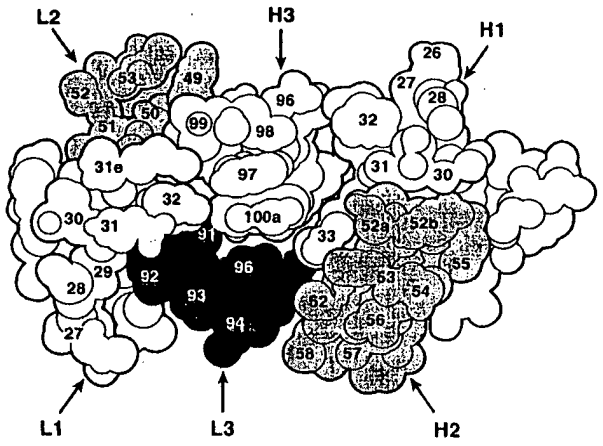


Fig. 1. The antigen binding site of an antibody. A space-filling model of the McPC603 antigen binding site depicting the superposition of five sections cut at 2 Å intervals is shown (Segal *et al.*, 1974). Residues 31a, 31b, 31c, 31d of the V_L domain and residue 52c of the V_H domain are above the plane of the paper. The structural repertoire available to the H1 and H2 regions of the human V_H domain has been previously described (Chothia *et al.*, 1992). Here we discuss the repertoire of main chain conformations available to the L1, L2 and L3 regions of the human V_κ domain.

Table 1. Genetic and structural defects in the human germline V_κ segments[a]

V_κ segment	Genetic defects	Structural defects[b]	
O14/O4		88G (C)	Loss of disulphide bridge and creation of an internal cavity
A4	Frame shift in V_κ exon (insertion of one nucleotide at residue 57)		
A21	Frame shift in V_κ exon (insertion of one nucleotide at residue 39)		
A7		23F (C)	Loss of disulphide bridge and small residue replaced by a larger one at an internal site
A29	No start codon at beginning of leader sequence (ATG replaced by ATA)		
A13	No start codon at beginning of leader sequence (ATG replaced by ATA)	37P (L/Q)	Loss of β-sheet hydrogen bonds by P
L22		5I (T) 22I (N/S/T) 24W (R) 56P (S/T) 76I (N/S)	A large number of mutations, at sites on the surface of the V_L domain, replace hydrophilic residues by hydrophobic residues
L10	Defective acceptor splice site in leader intron (AG replaced by GG)		
L13	No start codon at beginning of leader sequence (ATG replaced by GTG)	75F (I)	Increase in size of an absolutely conserved buried residue
		87C (F/Y)	Large reduction in size of a residue buried in V_L–V_H interface
B1	No start codon at beginning of leader sequence (ATG replaced by ATA)		

[a]Genetic and structural defects in the 51 V_κ segments which were described as 'potentially functional' or having 'minor defects' (Schäble and Zachau, 1993). References: O14/O4, Pargent *et al.* (1991); A4, Jaenichen *et al.* (1984); A21, Schäble *et al.* (1994); A7, A29, A13, Lautner-Rieske *et al.* (1992); L22, Pech *et al.* (1985); L10, Pech and Zachau (1994); L13, Huber *et al.* (1993); B1, Lorenz *et al.* (1988). For defects in the 25 'pseudogenes', see Schäble *et al.* (1994).
[b]Normal residues are given in parentheses.

sequences to determine how the assembly of V_κ and J_κ segments influences the canonical structure of the third antigen binding loop. In addition, the extent to which canonical structures are conserved during affinity maturation is determined. The implications of our results for understanding the mechanisms of immune recognition are then discussed.

The repertoire of human germline V_κ segments

For a germline V_κ segment to be functional, it must have (i) intact coding and regulatory sequences, (ii) functional recombination signal sequences (RSS) and (iii) encode a protein that can form a stable three-dimensional structure. On inspection of the sequences of the 76 V_κ segments, Schäble and Zachau (1993) described 25 as pseudogenes (see Schäble *et al.*, 1994).

Genetic and structural defects

The remaining 51 V_κ segments were described as being 'potentially functional' or having 'minor defects' (Schäble and Zachau, 1993). Segments with minor defects are those with one or two point mutations in the coding or regulatory regions that might render the segment non-functional, but which may exist as functional alleles in certain individuals (Schäble and Zachau, 1993). To examine the expression of each V_κ segment and its alleles, we compiled a large database of rearranged V_κ genes, derived from many individuals, containing 736 sequences (I.M.Tomlinson and J.P.L.Cox, in preparation), which extends previous studies of the expressed V_κ repertoire (Klein *et al.*, 1993; Cox *et al.*, 1994).

Eleven of the V_κ segments described as 'potentially functional' or having 'minor defects' have serious genetic and/or structural defects, such as the absence of start codons or mutations which are likely to seriously destabilize the three-dimensional structure of the protein (Table I); these segments are not expressed and are therefore classified as non-functional.

Functional V_κ segments

Thirty-eight of the remaining 40 V_κ segments are seen rearranged *in vivo*. Since the other two segments (A14 and L25) have no obvious defects which would prevent their functional expression, our failure to see them rearranged *in vivo* is probably due to them being located in the rarely used J_κ-distal portion of the V_κ locus (Cox *et al.*, 1994). The amino acid sequences of the 40 functional V_κ segments and their alleles are given in Table II [see Schäble and Zachau (1993) and Cox *et al.* (1994) for nucleotide sequences]. The functional segments are derived from six families: 19 V_κI, nine V_κII, seven V_κIII, one V_κIV, one V_κV and three V_κVI. Fifteen pairs of functional segments differ by <11 amino acids, some even encode identical gene products (Table II). Known alleles of the functional segments can alter the sequence of the V_κ exon, but only by up to three amino acids (see footnote to Table II). The introduction of stop codons in the V_κ exons of L4, L16 and A18, in the leader exon of O12, or the introduction of structural defects in O11, means that there may be as few as 35 functional segments in some individuals. The segments listed in Table II should therefore be regarded as the maximum number of functional V_κ segments in any one individual.

The use of these segments in the expressed repertoire

Table II. Amino acid translations of the human germline V_κ segments

	1 2 3 4 5 6 7 8 9 (0 0 0 abcdef 0 0 0 0 0 0 +++++)	CANONICAL STRUCTURE L1 L2 L3	SEQUENCE REFERENCES	
SUBGROUP I				
O12*/O2	DIQMTQSPSSLSASVGDRVTITCRASQSISS------YLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTP	2 1 (1)	O12/O2	a/a
O18/O8	DIQMTQSPSSLSASVGDRVTITCQASQDISN------YLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQYDNLP	2 1 (1)	O18/O8	b/b
A20	DIQMTQSPSSLSASVGDRVTITCRASQGISN------YLAWYQQKPGKVPKLLIYAASTLQSGVPSRFSGSGSGTDFTLTISSLQPEDVATYYCQKYNSAP	2 1 (X)	A20	c
A30	DIQMTQSPSSLSASVGDRVTITCRASQGIRN------DLGWYQQKPGKAPKRLIYAASSLQSGVPSRFSGSGSGTEFTLTISSLQPEDFATYYCLQHNSYP	2 1 (1)	A30	d
L14	NIQMTQSPSAMSASVGDRVTITCRARQGISN------YLAWFQQKPGKVPKHLIYAASSLQSGVPSRFSGSGSGTEFTLTISSLQPEDFATYYCLQHNSYP	2 1 (1)	L14	d
L1*	DIQMTQSPSSLSASVGDRVTITCRASQGISN------YLAWFQQKPGKAPKSLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYNSYP	2 1 (1)	L1	e
L15*	DIQMTQSPSSLSASVGDRVTITCRASQGISS------WLAWYQQKPEKAPKSLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYNSYP	2 1 (1)	L15	e
L4/18a*	AIQLTQSPSSLSASVGDRVTITCRASQGISS------ALAWYQQKPGKAPKLLIYDASSLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQFNSYP	2 1 (1)	L4/18a	f
L5/L19	DIQMTQSPSSVSASVGDRVTITCRASQGISS------WLAWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQANSFP	2 1 (1)	L5/L19	f/g
L8	DIQLTQSPSFLSASVGDRVTITCRASQGISS------YLAWYQQKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTEFTLTISSLQPEDFATYYCQQLNSYP	2 1 (1)	L8	f
L23	AIRMTQSPFSLSASVGDRVTITCWASQGISS------YLAWYQQKPAKAPKLFIYYASSLQSGVPSRFSGSGSGTDYTLTISSLQPEDFATYYCQQYYSTP	2 1 (1)	L23	d
L9	AIRMTQSPSSFSAS-GDRVTITCRASQGISS------YLAWYQQKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTDFTLTISCLQSEDFATYYCQQYYSYP	2 1 (1)	L9	f
L24	VIWMTQSPSLLSASTGDRVTISCRMSQGISS------YLAWYQQKPGKAPELLIYAASTLQSGVPSRFSGSGSGTDFTLTISCLQSEDFATYYCQQYYSFP	X 1 (1)	L24	h
L11	AIQMTQSPSSLSASVGDRVTITCRASQGIRN------DLGWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCLQDYNYP	2 1 (1)	L11	b
L12*	DIQMTQSPSTLSASVGDRVTITCRASQSISS------WLAWYQQKPGKAPKLLIYDASSLESGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYNSYS	2 1 (X)	L12	i
SUBGROUP II				
O11*/O1	DIVMTQTPLSLPVTPGEPASISCRSSQSLLDSDDGNTYLDWYLQKPGQSPQLLIYTLSYRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQRIEFP	3 1 (1)	O11/O1	a/a
A17	DVVMTQSPLSLPVTLGQPASISCRSSQSLVYS-DGNTYLNWFQQRPGQSPRRLIYKVSNRDSGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGTHWP	4 1 (1)	A17	c
A1	DVVMTQSPSLSLPVTLGQPASISCRSSQSLVYS-DGNTYLNWFQQRPGQSPRRLIYKVSNWDSGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGTHWP	4 1 (1)	A1	c
A18*	DIVMTQTPLSLSVTPGQPASISCKSSQSLLHS-DGKTYLYWYLQKPGQSPQLLIYEVSSRFSGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGIHLP	4 1 (1)	A18	c
A2	DIVMTQTPLSLSVTPGQPASISCKSSQSLLHS-DGKTYLYWYLQKPGQPPQLLIYEVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQSIQLP	4 1 (1)	A2	j
A19/A3	DIVMTQSPLSLPVTPGEPASISCRSSQSLLHS-NGYNYLDWYLQKPGQSPQLLIYLGSNRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQTP	4 1 (1)	A19/A3	c/k
A23	DIVMTQTPLSSPVTLGQPASISCRSSQSLVHS-DGNTYLSWLQQRPGQPPRLLIYKISNRFSGVPDRFSGSGAGTDFTLKISRVEAEDVGVYYCMQATQFP	4 1 (1)	A23	k
SUBGROUP III				
A27	EIVLTQSPGTLSLSPGERATLSCRASQSVSSS-----YLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYGSSP	6 1 (1)	A27	l
A11	EIVLTQSPATLSLSPGERATLSCGASQSVSSS-----YLAWYQQKPGLAPRLLIYDASSRATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYGSSP	6 1 (1)	A11	m
L2/L16*	EIVMTQSPATLSVSPGERATLSCRASQSVSS------NLAWYQQKPGQAPRLLIYGASTRATGIPARFSGSGSGTEFTLTISSLQSEDFAVYYCQQYNNWP	2 1 (1)	L2/L16	n/o
L6	EIVLTQSPATLSLSPGERATLSCRASQSVSS------YLAWYQQKPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQQRSNWP	2 1 (1)	L6	p
L20	EIVLTQSPATLSLSPGERATLSCRASQGVSS------YLAWYQQKPGQAPRLLIYDASNRATGIPARFSGSGPGTDFTLTISSLEPEDFAVYYCQQRSNWH	2 1 (X)	L20	q
L25	EIVMTQSPATLSLSPGERATLSCRASQSVSSS-----YLSWYQQKPGQAPRLLIYGASTRATGIPARFSGSGSGTDFTLTISSLQPEDFAVYYCQQDYNLP	6 1 (1)	L25	d
SUBGROUP IV				
B3	DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTRESGVPDRFSGSGSGTDFTLTISSLQAEDVAVYYCQQYYSTP	3 1 (1)	B3	q
SUBGROUP V				
B2	EIVLTQSPAFMSATPGDKVNISCKASQDIDD------DMNWYQQKPGEAAIFIIQEATTLVPGIPPRFSGSGYGTDFTLTINNIESEDAAYYFCLQHDNFP	2 1 (1)	B2	r
SUBGROUP VI				
A26/A10	EIVLTQSPDFQSVTPKEKVTITCRASQSIGS------SLHWYQQKPDQSPKLLIKYASQSFSGVPSRFSGSGSGTDFTLTINSLEAEDAATYYCHQSSSLP	2 1 (1)	A26/A10	c/s
A14	DVVMTQSPAFLSVTPGEKVTITCQASEGIGN------YLYWYQQKPDQAPKLLIKYASQSISGVPSRFSGSGSGTDFTFTISSLEAEDAATYYCQQGNKHP	2 1 (1)	A14	s

[a]V_κ segments marked with an asterisk have alleles that alter the amino acid sequence: O11a (ref. a) (35W→C 64G→D); O12a (ref. a) (10S→F 90Q→C 91S→G); for A18, the only known allele has a stop codon at position 88 (ref. c). Two of the rearranged counterparts of A18 have a Cys at position 88. These rearranged genes are probably derived from a functional allele of A18 which has a Cys at position 88; LFVK431 is an allele of L1 (ref. t) (61R→K); L15a (ref. i) (26S→R); the allele L4/18a is functional, but has not been unambiguously mapped to L4 and/or L18 (Schäble and Zachau, 1993). Mapped alleles are L4 (ref. f) (35W→* 93S→N) and L18 (ref. g) (93S→N); L16c (ref. n) (5T→M 94W→*); L12a (ref. d) (50D→K).

First publication of full-length sequence: (a) Pargent *et al.* (1991); (b) Scott *et al.* (1991); (c) Lautner-Rieske *et al.* (1992); (d) Huber *et al.* (1993); (e) Bentley and Rabbitts (1983); (f) Pech *et al.* (1984); (g) Pech *et al.* (1985); (h) Jaenichen *et al.* (1984); (i) Bentley and Rabbitts (1980); (j) Scott *et al.* (1989); (k) Staubinger *et al.* (1988a); (l) Radoux *et al.* (1986); (m) Chen *et al.* (1986); (n) Liu *et al.* (1989); (o) Chen *et al.* (1987); (p) Pech and Zachau (1984); (q) Klobeck *et al.* (1985); (r) Stavnezer *et al.* (1985); (s) Staubinger *et al.* (1988b); (t) Foroni (1990).

is highly biased. Not only is there a strong bias towards use of V_κ segments from the J_κ-proximal region of the locus (97% of the expressed repertoire is derived from V_κ segments in the J_κ-proximal region), but there is also a bias towards the use of particular V_κ segments. For example, the single segment A27 accounts for almost 30% of the expressed repertoire, whereas the 22 least used V_κ segments account for only 6% of the expressed repertoire. Those less frequently used segments include several with defects in the promoter region (L8, L9) or heptamer RSS (A20, A26, A10, A14, L4, L18) which are likely to reduce the efficiency of transcription or rearrangement (Akamatsu *et al.*, 1994; Stiernholm and Berinstein, 1995).

Canonical structures of the first antigen binding loop of the V_κ domain

Sequence variation of residues 24–34 in the V_L domain led Kabat and Wu (1971) to predict that this region is involved in antigen binding and to call it the first complementarity determining region (CDR1). Later, an analysis of the three-dimensional structures showed that residues 24–25 and 33–34 are part of the β-sheet framework: sequence changes at these positions do not alter the main chain conformation (Chothia and Lesk, 1987). The region outside the β-sheet framework, 26–32, is called L1.

Further analysis suggested that nearly all L1 regions in V_κ domains have one of six canonical structures (Chothia and Lesk, 1987; Chothia *et al.*, 1989; unpublished data). All have residues 26, 27, 28, 29 and 32 packed against the framework in the same extended conformation (Figure 2). They differ in the number of residues between positions 29 and 32 (Figure 2). All six are found in murine

Table III.

L1 Canonical structure	Length of L1 region (26–32 inclusive) and the residues between 29 and 32
2	7: 30, 31
6	8: 30, 31, 31a
4	12: 30, 31, 31a, 31c, 31d, 31e, 31f
3	13: 30, 31, 31a, 31b, 31c, 31d, 31e, 31f

Table IV. Residues at key sites for the canonical structures of the V_κ domain

(A) L1 region

Canonical structure	Number of residues in L1 region (26–32 inclusive)	Sequence characteristics of key residues in known structures						Number of germline segments with this canonical structure
		2	25	29		33	71	
L1: 2	7	I: 18	A: 18	I: 16 V: 2		I: 1 L: 15 V: 2	F: 7 Y: 11	26
		2	25	29		33	71	
L1: 6	8	N: 1[a]	A: 1	V: 1		L: 1	Y: 1	3
		2	25	29	31d	33	71	
L1: 4	12	L: 1 V: 9	S: 9 P: 1	I: 3 L: 7	G: 10	F: 1 L: 9	F: 10	7
		2	25	29	31d	33	71	
L1: 3	13	I: 5	S: 5	L: 4 V: 1	E: 1 Q: 3 S: 1	L: 5	F: 5	3

[a]For L1 canonical structure 6, there is only a single crystal structure (the murine antibody 1F7; Haynes *et al.*, 1994). Although the Asn at position 2 in this structure (see Figure 2) is unusual (it is normally a medium-sized hydrophobic residue, either Ile, Leu or Val), one-third of murine V_κVI sequences also contain Asn at this position (Kabat *et al.*, 1991). However, in the 1F7 structure it does not seem to affect the main chain conformation of residues 26–29, which is conserved in all known L1 structures. We would expect Thr 2 in the human V_κ segment B2 to be accomodated in a similar manner.

(B) L2 region

Canonical structure	Number of residues in L2 region (50–52 inclusive)	Sequence characteristics of key residues in known structures		Number of germline segments with this canonical structure
		48	64	
L2: 1	3	I: 39 V: 2	G: 41	40

(C) L3 region

Canonical structure	Number of residues in L3 region (91–96 inclusive)	Sequence characteristics of key residues in known structures								Number of rearranged sequences with this canonical structure
		90	91	92	93	94	95	96	97	
L3: 1	6	H: 2 N: 5 Q: 30	–	–	–	–	P: 37	–	S: 1 T: 36	324
		90	91	92	93	94	95	96	97	
L3: 2	6	Q: 1	–	–	–	P: 1	–	–	T: 1	1
		90	91	92	93	94	95		97	
L3: 3	5	Q: 1	–	–	–	–	–		–	49
		90	91	92	93	94			97	
L3: 4	4	Q: 1	–	–	–	–			S: 1	1
		90	91	92	93	94	95	96	97	
L3: 5	7	Q: 1	–	–	–	–	–	P: 1	T: 1	67

antibodies, but only four of the six are present in humans. In humans, the four structures are listed in Table III.

The canonical structure numbers are historical in origin. Mice also encode canonical structure 1, that has one residue between 29 and 32, and canonical structure 5, with six residues between 29 and 32. The non-functional human V_κ segment B1 has the sequence characteristics of canonical structure 5 (Table I).

The key residues that determine the conformation of the known L1 canonical structures for the V_κ domain are at positions 2, 25, 29, 33 and 71, which are involved in the packing of L1 against the framework (Figure 2). In canonical structures 3 and 4, residue 31d has positive ϕ,ψ values which produce some steric strain in residues other than Gly, Asn and Asp. The conformation is facilitated by the presence of these residues, but they are not essential.

There are now several crystal structures available for each of the L1 canonical structures 2, 3 and 4, and a single structure available for canonical structure 6. The residues at the key sites in these structures are given in Table IVA.

The L1 structures implicit in the germline repertoire

Thirty-nine human V_κ segments have L1 regions that correspond in size and key residues to one of four canonical structures (Tables II and IVA). The segment L24 (subgroup I) has all the characteristics of canonical structure 2, except that the key residue at position 25 is Met, rather than Ala or Ser. It is difficult to determine whether the large Met side chain at this site would produce a strained protein with canonical structure 2 or a change in conformation to form a new canonical structure.

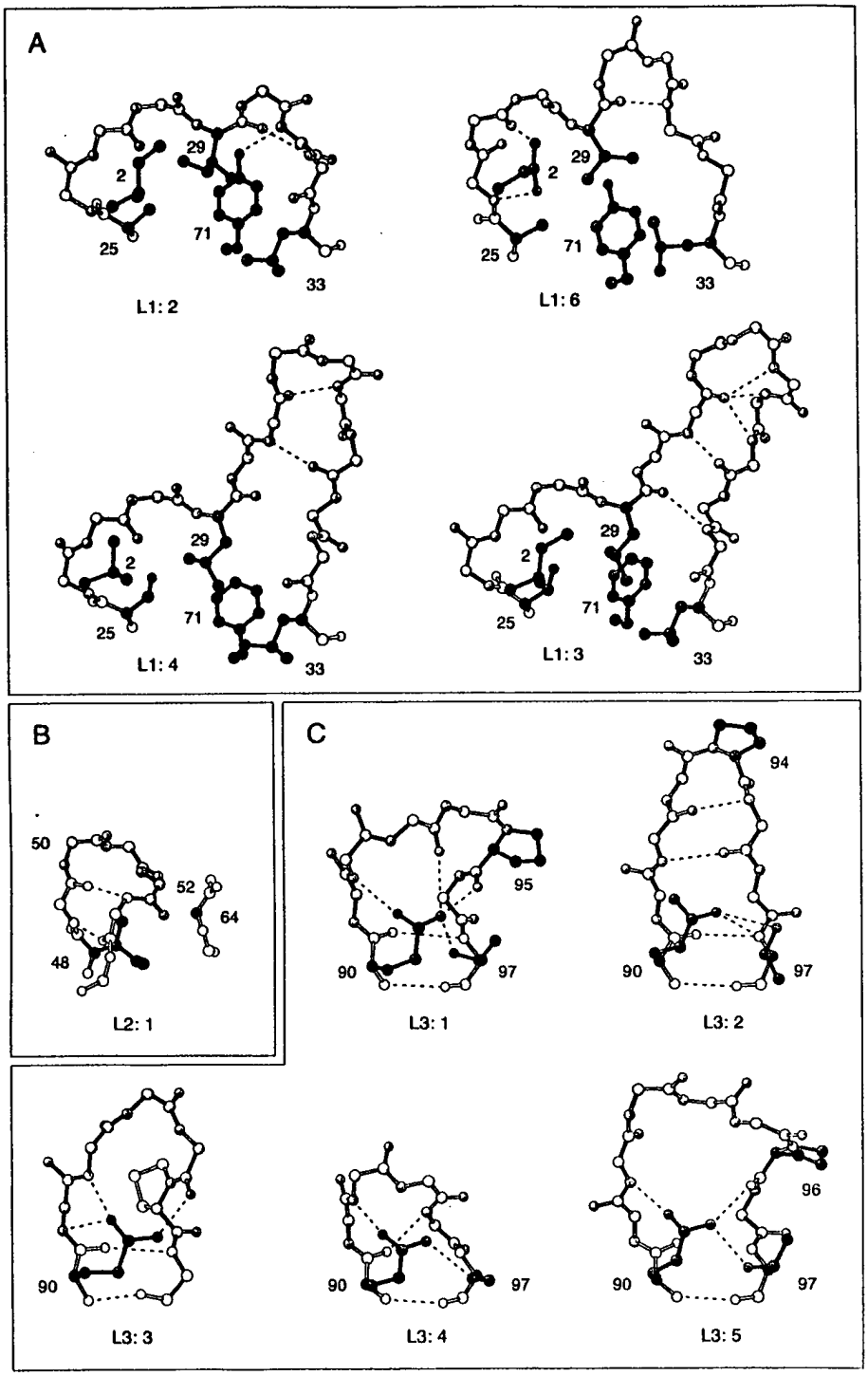


Fig. 2. Canonical structures of the antigen binding loops encoded by human V_K domains. Each drawing is labelled La:b, where La corresponds to the antigen binding loop, L1 (A), L2 (B) or L3 (C), and b is the number of the canonical structure. In each case, we show the main chain of the antigen binding loops (yellow) and the Cα atom/side chain of the key residues which are the main determinants of its conformation (red). In addition, the conformations of L1: 3 and L1: 4 are facilitated by the presence of Gly, Asn and Asp at 31d which have positive ϕ,ψ values (see the text). The conformations shown here are taken from antibodies whose structures have been determined by X-ray crystallography: L1: 2 from H52 (Eigenbrot *et al.*, 1994); L1: 6 from 1F7 (Haynes *et al.*, 1994); L1: 4 from 4-4-20 (Whitlow, 1995); L1: 3 from 17/9 (Rini *et al.*, 1992); L2: 1 from KOL (Marquart *et al.*, 1980); L3: 1 from REI (Epp *et al.*, 1975); L3: 2 from J539 (Suh *et al.*, 1986); L3: 3 from HyHEL5 (Sheriff *et al.*, 1987); L3: 4 from 3D6 (He *et al.*, 1992); L3: 5 from AN02 (Brünger *et al.*, 1991).

The L1 structures implicit in the expressed repertoire

Our database of rearranged V_K sequences has 736 entries, although some of them are incomplete. Of these, 401 (55%) are derived from germline V_K segments which encode canonical structure 2 for the L1 region, 54 (7%) from segments which encode canonical structure 3, 65 (9%) from segments which encode canonical structure 4

Table V. Somatic mutation at the key sites for the L1 and L2 regions in the human V_κ domain

Region	Key residue	Number of sequences containing key residue	Number of sequences with no change in key residue	Conservative changes and their occurrences	Non-conservative changes and their occurrences
L1	2	333	254	I→F: 2 I→L: 67 I→T: 4 I→V: 5 V→I: 1 Total: 79[a]	
	25	715	686	A→G: 1 A→S: 9 A→T: 12 S→A: 1 S→T: 1 Total: 24	A→E: 2 A→V: 2 S→L: 1 Total: 5
	29	714	630	I→L: 1 I→M: 1 I→V: 11 V→I: 44 V→L: 18 V→M: 1 V→T: 1 Total: 77	I→K: 1 V→A: 2 V→F: 3 V→G: 1 Total: 7
	31d	115	108	N→D: 1 Total: 1	G→E: 1 G→R: 2 N→K: 2 N→S: 1 Total: 6
	33	717	701	L→F: 1 L→V: 11 Total: 12	L→S: 4 Total: 4
	71	730	723	F→Y: 3 Total: 3	F→L: 3 Y→C: 1 Total: 4
L2	48	723	709	I→L: 1 I→M: 5 I→V: 4 Total: 10	I→F: 3 I→Y:1 Total: 4
	64	722	713	G→A: 7 Total: 7	G→S: 1 G→V: 1 Total: 2

[a]Most of these changes are probably introduced by PCR primers at the beginning of framework 1.

and 216 (29%) from segments which encode canonical structure 6.

We examined their sequences to determine whether they have insertions, deletions or mutations that produce L1 conformations different to those encoded by the germline segments. None has insertions or deletions, relative to the corresponding germline segments. The number of sequences in which changes have occurred at the key sites, and the nature of these changes, are given in Table V. The changes are classified as conservative or non-conservative. Conservative changes are those that are unlikely to make any large changes in the conformation of the loop, but which may introduce small distortions or slightly shift its mean position and the envelope of alternative low-energy positions available to it. Non-conservative changes are likely to make a major alteration in the conformation or lower its stability. At four sites, 25, 31d, 33 and 71, only a few sequences contain changes. Two other sites, 2 and 29, are more frequently changed. Although the majority of changes at residue 2 are likely to have been introduced by primers based at the beginning of framework 1, those that occur at the other key residues probably represent genuine somatic mutations. The majority of these changes are conservative. Thus, the key sites for L1 are unmutated, or undergo only conservative changes, in >96% of the rearranged sequences and undergo non-conservative changes in <4% of the rearranged sequences.

Canonical structure of the second antigen binding loop of the V_κ domain

Sequence variation of residues 50–56 in V_L domains led Kabat and Wu (1971) to predict that this region is involved in antigen binding and to call it the second complementarity determining region (CDR2). Later, an analysis of the three-dimensional structures showed that residues 53–56 are part of the β-sheet framework structure: sequence changes at these positions do not alter the main chain conformation (Chothia and Lesk, 1987). The region outside the β-sheet framework, 50–52, is called L2.

Further analysis indicated that L2 regions in V_κ domains have a single canonical structure defined by there being a single residue between 50 and 52, Ile or Val at 48 and Gly at 64 (Chothia and Lesk, 1987; Steipe *et al.*, 1992) (Figure 2 and Table IVB).

The L2 structure implicit in the germline and expressed repertoires

All 40 functional germline V_κ segments have residues Ile at site 48, Gly at site 64 and one residue between residues 50 and 52 (Tables II and IVB); this means they all have the features of the single canonical structure known for the L2 region. Consequently, all 736 rearranged sequences are derived from germline V_κ segments with this canonical structure. None of these has insertions or deletions in the L2 region, relative to the corresponding germline segments, and only six have non-conservative changes at L2

Table VI.

J_κ segment	Sequence[a]
1	W T F G Q G T K V E I K R
2	Y T F G Q G T K L E I K R
3	F T F G P G T K V D I K R
4	L T F G G G T K V E I K R
5	I T F G Q G T R L E I K R

[a]First residue is position 96, last residue is position 108.

key sites (Table V). Thus, the key sites for L2 are unmutated or undergo conservative changes in >99% of the rearranged sequences and undergo non-conservative changes in <1% of the rearranged sequences.

Canonical structures of the third antigen binding loop of the V_κ domain

Sequence variation of residues 89–97 in V_L domains led Kabat and Wu (1971) to predict that this region is involved in antigen binding and to call it the third complementarity determining region (CDR3). Later, an analysis of the three-dimensional structures showed that residues 89–90 and 97 are part of the β-sheet framework structure (Chothia and Lesk, 1987). The region outside the β-sheet framework, 91–96, is called L3.

The L3 region is formed by the joining of the V_κ and J_κ segments. The joining process involves trimming and repair of the V and J segments, as well as nucleotide insertions (Milstein.*et al.*, 1992). This produces variations in both the length of the L3 region and in the identity of residues at the join. The structure of the L3 region is therefore determined not only by the sequence of the germline V_κ and J_κ segments, but also by the way in which they are assembled. There are five J_κ segments, all functional and with the sequences listed in Table VI (Hieter *et al.*, 1982).

The L3 canonical structures known at present

The analysis of the three-dimensional structures of V_κ L3 regions has so far identified five canonical structures (Chothia and Lesk, 1987; Chothia *et al.*, 1989; Brünger *et al.*, 1991; He *et al.*, 1992). The five structures are illustrated in Figure 2 and in Table IVC we describe their size and sequence characteristics.

Canonical structure 1 has six residues, a Gln, Asn or His at position 90, a *cis*-Pro at 95 and a Ser or Thr at 97. Thirty-seven of the 40 functional V_κ segments have Gln at position 90 and Pro at position 95 (Table II). All five J_κ segments have Thr at 97 (Table VI). Therefore, these 37 V_κ segments can form canonical structure 1 if the joining process juxtaposes the last codon of the V_κ segment with the first codon of the J_κ segment. The two V_κ segments that have Ser and His at position 95 (Table II) can also produce this canonical structure if V_κ–J_κ joining modifies this codon to produce a Pro at position 95.

All the other known canonical structures are produced by modification of the germline segments by the joining process. Canonical structure 2 has six residues, but the Pro is at position 94 rather than 95. Since none of the functional V_κ segments has a Pro at position 94 and V_κ–J_κ joining events that modify this codon are infrequent (I.M.Tomlinson and J.P.L.Cox, in preparation), this conformation is likely to be rare in human V_κ domains. In canonical structure 3, the joining process has deleted one residue, leaving only five (Chothia *et al.*, 1989). In canonical structure 4, the joining process has deleted two residues leaving only four (He *et al.*, 1992). In canonical structure 5, the joining process adds a residue to produce a seven-residue L3 region.

Table VII.

Number of residues in the L3 region, 91–96	Number of rearranged sequences
1	1
2	2
3	2
4	1
5	54
6	427
7	119
8	26
9	1

The L3 structures implicit in the expressed repertoire

Of the 736 rearranged V_κ sequences in our database, 633 contain complete L3 regions without frame shifts. Junctional diversity has produced some variations in the length of L3, but it is limited: in 99% of the rearranged sequences the number of residues in the L3 region, 91–96, is between five and eight (Table VII).

We examined the L3 regions to determine which have the characteristics of one of the known canonical structures (Table IVC). Of those with six residues, three-quarters have the sequence characteristics of canonical structure 1. Most of those with five residues have the characteristics of canonical structure 3 and most of those with seven residues have the characteristics of canonical structure 5. Canonical structures 2 and 4 are rare, each occurring in only one sequence. In total, 70% of the rearranged sequences have the characteristics of one of the five known canonical structures (Table IVC).

The sequences of the remaining 191 rearranged genes do not correspond to a known L3 canonical structure. However, it is likely that most of these will have one of a small number of conformations. For example, we would expect most of the six-residue regions that do not have a Pro at positions 94 or 95 to have a simple hairpin structure with a conventional two-residue turn at its tip. A total of four or five canonical structures (including three for which crystal structures are available) probably describe the conformations of >90% of human L3 regions.

Because it is difficult to distinguish between residue changes introduced during V_κ–J_κ joining and those introduced by somatic mutation at key residues 95 and 97, we cannot exclude the possibility that the canonical structure of L3 changes during affinity maturation due to mutation at these positions. However, of the 729 rearranged sequences that include key residue 90, 90% are unmutated (657 sequences), 5% undergo conservative changes (Gln changes to His in 38 sequences), 4% undergo non-conservative changes (29 sequences with Gln to residues other than Asn or His) and in 1% the nature of the changes

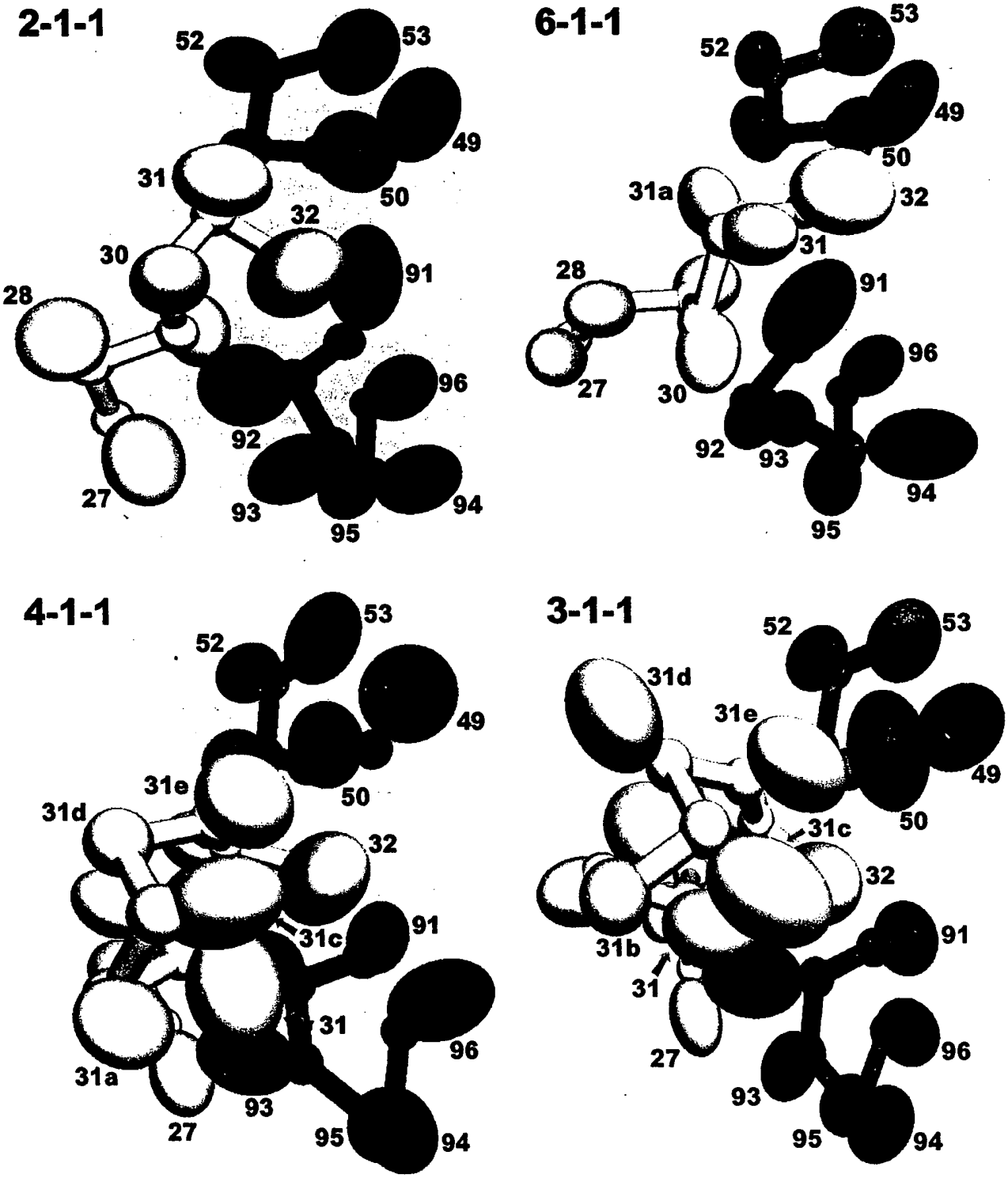


Fig. 3. Structures of the common antigen binding sites in human V_K domains. Residues are represented by their Cα atoms and ellipsoids for the sidechains. The numbers on the top left of each drawing correspond to the canonical structure of the three antigen binding loops (L1, L2 and L3). More than 96% of the human V_K domains have canonical structures 2, 3, 4 or 6 for the first antigen binding loop (L1) and canonical structure 1 for the second antigen binding loop (L2). Approximately half of the human V_K domains have canonical structure 1 for the third antigen binding loop (L3).

is uncertain (five sequences where Lys changes to Gln or Asn). This suggests that L3 canonical structures are also highly conserved during affinity maturation.

Discussion

The conformations of canonical structures

Canonical structures describe the common folds of the antigen binding loops of antibodies. When they occur in different antibodies how similar are their conformations? A number of different antibody structures with the same canonical structures have been determined accurately at high resolution, including several in complex with their antigens. Analysis of these structures shows that the local conformation of the canonical structures is highly conserved (A.M.Lesk and C.Chothia, in preparation). Normally, those that are small or medium sized (e.g. L2 and the smaller L1 regions) have rms differences in atomic position of 0.5 Å or less. For those that are large (e.g. the larger L1 and L3 regions) the values are 0.8 Å or less. Very occasionally, hydrogen bonding from within the variable domain or by the antigen rotates a peptide group into an unusual orientation such that the canonical structure has rms differences of ~1 Å when compared with those of normal conformations.

Sequence differences in the framework regions do not change the local conformation of canonical structures, but can shift their relative positions by up to ~1.3 Å (Chothia and Lesk, 1987).

The interface between V_L and V_H domains has a large central region formed by conserved framework residues and a peripheral region formed by residues from the antigen binding loops. The conserved residues fix the overall packing of the V_L and V_H domains. However, differences in the antigen binding loops, and the contacts formed by antigens that penetrate the contact area, modulate the detailed geometry of the packing such that in extreme cases the relative orientation of V_L and V_H domains can differ by up to ~15° (Colman *et al.*, 1987; Lascombe *et al.*, 1989; Stanfield *et al.*, 1993). These differences in orientation shift the relative positions of the canonical structures in the two domains by 1–2 Å in most cases and by 3 Å in a few.

Mechanisms of immune recognition

The function of antibodies is to bind antigens with high specificity and affinity. In antigen–antibody complexes, binding sites for protein antigens occupy ~800 $Å^2$, which is close to one-third of the total surface presented to the solvent by the antigen binding loops. The sites that bind peptides and haptens are somewhat smaller [see Davies *et al.* (1990) and Wilson and Stanfield (1993, 1994) for reviews]. Examination of accurate high-resolution complexes demonstrates that the antigen–antibody interface is close packed, internal cavities are rare, and hydrogen bonding, van der Waals contacts and solvent molecules combine to make the packing density like that in crystals. These features are the same as those found in the recognition sites of non-immunoglobulin protein–protein complexes (Janin and Chothia, 1990).

The purpose of the genetic mechanisms that produce high-affinity antibodies is therefore to create structures that either have complementary surfaces of sufficient size to bind antigens or can produce such surfaces by low-energy conformational changes. Isolated loop structures, such as those that occur in certain H3 regions with parts free from tertiary constraints, can switch between different conformations (Wilson and Stanfield, 1994) but, in general, the intrinsic flexibility of antigen binding loops is limited with only a small envelope of alternative positions (Bhat *et al.*, 1990; Gerstein *et al.*, 1994).

The precise structural requirements for high-affinity antibodies are achieved in two stages. The assembly of germline segments produces a range of antibodies that are believed to be sufficiently diverse to recognize almost all antigens with at least moderate affinity. The small repertoire of canonical structures (Figure 3) provides a framework on which to display a highly diverse set of residues for selection during the primary immune response. Indeed, the residues that have the highest germline variability are clustered at the centre of the binding site (Chothia *et al.*, 1992; I.M.Tomlinson and J.P.L.Cox, in preparation).

Subsequent somatic mutation and selection leads to an improvement in the affinity of the antigen–antibody interaction. Typically, this maturation process increases affinity constants by a factor of 10^2–10^3 (Berek and Milstein, 1987). We have shown that during affinity maturation, human antibodies conserve the canonical loop structures selected during the primary response. There are no examples of somatic insertions or deletions in the L1 and L2 regions. Together, the key sites for the L1 and L2 regions have undergone non-conservative mutation in <4% of the rearranged sequences. Although it is more difficult to demonstrate conservation of the L3 canonical structure during affinity maturation (see above), in the vast majority of rearranged sequences the key site at position 90 is unmutated or undergoes conservative mutation. Our analysis of the human V_H repertoire demonstrated a similar conservation for the H1 and H2 regions (Chothia *et al.*, 1992). Although there are no data available for somatic mutation in human H3 regions, data have been obtained in mice by sequencing anti-hapten antibodies during an ongoing immune response. Reviewing these studies, Berek and Milstein (1987) noted that the length and certain residues of the H3 region are not modified by somatic mutation.

Conservative mutations engineered into residues that make direct contact with the antigen can change the affinity of the antibody–antigen interaction by factors of up to 10^3, whilst conservative mutations engineered into the key residues that contact the antigen binding loops can change the affinity by factors of 3–10 (Amit *et al.*, 1986; Alzari *et al.*, 1990; Sharon, 1990; Denzin *et al.*, 1991; Foote and Winter, 1992; Hawkins *et al.*, 1992, 1993; Schildbach *et al.*, 1993; Chacko *et al.*, 1995; Chen *et al.*, 1995). Here, we have shown that although non-conservative mutations at key residues that contact the antigen binding loops are rarely observed, >17% of rearranged sequences have undergone conservative mutation at these sites. Whilst these changes are unlikely to disrupt or change the canonical structures of the antigen binding loops (see above), they may well play a role in affinity maturation by producing small shifts in their geometry.

References

Akamatsu,Y., Tsurushita,N., Nagawa,F., Matsuoka,M., Okazaki,K., Imai,M. and Sakano,H. (1994) Essential residues in V(D)J recombination signals. *J. Immunol.*, **153**, 4520–4529.

Amit,A.G., Mariuzza,R.A., Phillips,S.E.V. and Poljak,R.J. (1986) Three dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science*, **233**, 747–753.

Alzari,P.M., Spinelli,S., Mariuzza,R.A., Boulot,G., Poljak,R.J., Jarvis,J.M. and Milstein,C. (1990) Three-dimensional structure determination of an anti-2-phenyloxazolone antibody: the role of somatic mutation and heavy/light chain pairing in the maturation of an immune response. *EMBO J.*, **9**, 3807–3814.

Bentley,D.L. and Rabbitts,T.H. (1980) Human immunoglobulin variable region genes—DNA sequences of two V_κ genes and a pseudogene. *Nature*, **288**, 730–733.

Bentley,D.L. and Rabbitts,T.H. (1983) Evolution of immunoglobulin V genes: Evidence indicating that recently duplicated human V_κ sequences have diverged by gene conversion. *Cell*, **32**, 181–189.

Berek,C. and Milstein,C. (1987) Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.*, **96**, 23–41.

Bhat,T.N., Bentley,G,A., Fischmann, Boulot,G. and Poljak,R.J. (1990) Small rearrangements in structures of Fv and Fab fragments of an antibody D1.3 on antigen binding. *Nature*, **347**, 483–485.

Brünger,A.T., Leahy,D.J., Hynes,T.R. and Fox,R.O. (1991) 2.9 Å resolution structure of an anti-dinitrophenyl-spin-label monoclonal antibody Fab fragment with bound hapten. *J. Mol. Biol.*, **221**, 239–256.

Chacko,S., Silverton,E., Kam-Morgan,L., Sandra-Gill,S., Cohen,G. and Davies,D. (1995) Structure of an antibody-lysozyme complex: unexpected effect of a conservative mutation. *J. Mol. Biol.*, **245**, 261–274.

Chen,C., Roberts,V.A., Stevens,S., Brown,M., Stenzel-Poore,M.P. and Rittenberg,M.B. (1995) Enhancement and destruction of antibody function by somatic mutation: unequal occurrence is controlled by V gene combinatorial associations. *EMBO J.*, **14**, 2784–2794.

Chen,P.P., Albrandt,K., Orida,N.K., Radoux,V., Chen,E.Y., Schrantz,R., Liu,F.-T. and Carson,D.A. (1986) Genetic basis for the cross-reactive idiotypes on the light chains of human IgM anti-IgG antibodies. *Proc. Natl Acad. Sci. USA*, **83**, 8318–8322.

Chen,P.P., Robbins,D.L., Jirik,F.R., Kipps,T.J. and Carson,D.A. (1987) Isolation and characterisation of a light chain variable region gene for human rheumatoid factors. *J. Exp. Med.*, **166**, 1900–1905.

Chothia,C. and Lesk,A.M. (1987) Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.*, **196**, 901–917.

Chothia,C. *et al.* (1989) Conformations of immunoglobulin hypervariable regions. *Nature*, **342**, 877–883.

Chothia,C., Lesk,A.M., Gherardi,E., Tomlinson,I.M., Walter,G., Marks,J.D., Llewelyn,M.B. and Winter,G. (1992) Structural repertoire of the human V_H segments. *J. Mol. Biol.*, **227**, 799–817.

Colman,P.M., Laver,W.G., Varghese,J.N., Baker,A.T., Tulloch,P.A., Air,G.M. and Webster,R.G. (1987) Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature*, **326**, 358–363.

Cox,J.P.L., Tomlinson,I.M. and Winter,G. (1994) A directory of human germ-line V_κ segments reveals a strong bias in their usage. *Eur. J. Immunol.*, **24**, 827–836.

Davies,D.R., Padlan,E.A. and Sheriff,S. (1990) Antibody-antigen complexes. *Annu. Rev. Biochem.*, **59**, 439–473.

Denzin,L.K., Whitlow,M. and Voss,E.W. (1991) Single-chain site-specific mutations of fluorescein-aminoacid contact residues in high affinity monoclonal antibody 4-4-20. *J. Biol. Chem.*, **266**, 14095–14103.

Eigenbrot,C., Gonzalez,T., Mayeda,J., Carter,P., Werther,W., Hotaling,T., Fox,J. and Kessler,J. (1994) X-ray structures of fragments from binding and non-binding versions of a humanized anti-CD18 antibody: structural indications of the key role of V_H residues 59 to 65. *Proteins*, **18**, 49–62.

Epp,O., Colman,P., Fehlhammer,H., Bode,W., Schiffer,M. and Huber,R. (1975) Crystal and molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI. *Eur. J. Biochem.*, **45**, 513–524.

Foote,J. and Winter,G. (1992) Antibody framework residues affecting the conformation of the hypervariable loops. *J. Mol. Biol.*, **224**, 487–499.

Foroni,L. (1990) PhD Thesis, University of London, UK.

Gerstein,M., Lesk,A.M. and Chothia,C. (1994) Structural mechanisms for domain movements in proteins. *Biochemistry*, **33**, 6739–6749.

Hawkins,R.E., Russell,S.J. and Winter,G. (1992) Selection of phage antibodies by binding affinity: Mimicking affinity maturation. *J. Mol. Biol.*, **226**, 889–896.

Hawkins,R.E., Russell,S.J., Baier,M. and Winter,G. (1993) The contribution of contact and non-contact residues of antibody in the affinity of binding to antigen: The interaction of mutant D1.3 antibodies with lysozyme. *J. Mol. Biol.*, **234**, 958–964.

Haynes,M.R., Stura, E.A., Hilvert,D. and Wilson,I.A. (1994) Routes to catalysis: structure of a catalytic antibody and comparison with its natural counterpart. *Science*, **263**, 646–652.

He,X.M., Ruker,F., Casale,E. and Carter,D.C. (1992) Structure of a human monoclonal antibody Fab fragment against gp41 of human immunodeficiency virus type 1. *Proc. Natl Acad. Sci. USA*, **89**, 7154–7158.

Hieter,P.A., Maizel,J.V. and Leder,P. (1982) Evolution of human immunoglobulin κ J region genes. *J. Biol. Chem.*, **257**, 1516–1522.

Huber,C., Schäble,K.F., Huber,E., Klein,R., Meindl,A., Thiebe,R., Lamm,R. and Zachau,H.G. (1993) The V_κ genes of the L regions and the repertoire of V_κ gene sequences in the human germ line. *Eur. J. Immunol.*, **23**, 2868–2875.

Jaenichen,H.-R., Pech,M., Lindenmaier,W., Wildgruber,N. and Zachau,H.G. (1984) Composite human V_κ genes and a model of their evolution. *Nucleic Acids Res.*, **12**, 5249–5263.

Janin,J. and Chothia,C. (1990) The structure of protein–protein recognition sites. *J. Biol. Chem.*, **265**, 16027–16030.

Kabat,E.A. and Wu,T.T. (1971) Attempts to locate complementarity determining residues in the variable positions of light and heavy chains. *Ann. NY Acad. Sci.*, **190**, 382–393.

Kabat,E.A., Wu,T.T., Perry,H.M., Gottesman,K.S. and Foeller,C. (1991) *Sequences of Proteins of Immunological Interest*, 5th edn. Public Health Service, NIH, Washington, DC.

Klein,R., Jaenichen,R. and Zachau,H.G. (1993) Expressed human immunoglobulin kappa genes and thier hypermutation. *Eur. J. Immunol.*, **23**, 3248–3271.

Klobeck,H.-G., Bornkamm,G.W., Combriato,G., Mocikat,R., Pohlenz,H.-D. and Zachau,H.G. (1985) Subgroup IV of human immunoglobulin K light chains is encoded by a single germline gene. *Nucleic Acids Res.*, **13**, 6515–6529.

Lascombe,M.-B. *et al.* (1989) Three dimensional structure of Fab R19.9: a monoclonal murine antibody specific for the p-azobenzene-arsonate group. *Proc. Natl Acad. Sci. USA*, **86**, 607–611.

Lautner-Rieske,A., Huber,C., Meindl,A., Pargent,W., Schäble,K.F., Thiebe,R., Zocher,I. and Zachau,H.G. (1992) The human immunoglobulin κ locus. Characterization of the duplicated A regions. *Eur. J. Immunol.*, **22**, 1023–1029.

Liu,M.-F., Robbins,D.L., Crowley,J.J., Sinha,S., Kozin,F., Kipps,T.J., Carson,D.A. and Chen,P.P. (1989) Characterization of four homologous L chain variable region genes that are related to 6B6.6 idiotype positive human rheumatoid factor L chains. *J. Immunol.*, **142**, 688–694.

Lorenz,W., Schäble,K.F., Thiebe,R., Stavnezer,J. and Zachau,H.G. (1988) The J_κ proximal region of the human K locus contains three uncommon Vκ genes which are arranged in opposite transcriptional polarities. *Mol. Immunol.*, **25**, 479–484.

Marquart,M., Deisenhofer,J., Huber,R. and Palm,W. (1980) Crystallographic refinement and atomic models of the intact immunoglobulin molecule Kol and its antigen-binding fragment at 3.0 Å and 1.9 Å resolution. *J. Mol. Biol.*, **141**, 369–391.

Milstein,C., Even,J., Jarvis,J.M., Gonzalez-Fernandez,A. and Gherardi,E. (1992) Non-random features of the repertoire expressed by the members of one V_κ gene family and of the V-J recombination. *Eur. J. Immunol.*, **22**, 1627–1634.

Pargent,W., Meindl,A., Thiebe,R., Mitzel,S. and Zachau,H.G. (1991) The human immunoglobulin κ locus. Characterization of the duplicated O regions. *Eur. J. Immunol.*, **21**, 1821–1827.

Pech,M. and Zachau,H.G. (1984) Immunoglobulin genes of different subgroups are interdigitated within the V_κ locus. *Nucleic Acids Res.*, **12**, 9229–9236.

Pech,M., Jaenichen,H.-R., Pohlenz,H.-D., Neumaier,P.S., Klobeck,H.-G. and Zachau,H.G. (1984) Organization and evolution of a gene cluster for human immunoglobulin variable regions of the kappa type. *J. Mol. Biol.*, **176**, 189–204.

Pech,M., Smola,H., Pohlenz,H.-D., Staubinger,B., Gerl,R. and Zachau,H.G. (1985) A large section of the gene locus encoding human immunoglobulin variable regions of the kappa type is duplicated. *J. Mol. Biol.*, **183**, 291–299.

Radoux,V., Chen,P.P., Sorge,J.A. and Carson,D.A. (1986) A conserved

human germline V_κ gene directly encodes rheumatoid factor light chains. *J. Exp. Med.*, **164**, 2119–2124.
Rini,J.M., Schulze-Gahmen,U. and Wilson,I. (1992) Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science*, **255**, 959–965.
Schäble,K.F. and Zachau,H.G. (1993) The variable genes of the human immunoglobulin κ locus. *Biol. Chem. Hoppe-Seyler*, **374**, 1001–1022.
Schäble,H.F., Thiebe,R., Flügel,A., Meindl,A. and Zachau,H.G. (1994) The human immunoglobulin κ locus: Pseudogenes, unique and repetitive sequences. *Biol. Chem. Hoppe-Seyler*, **375**, 189–199.
Schildbach,J.F., Near,R.I., Bruccoleri,R.E., Haber,E., Jeffrey,P.D., Novotny,J., Sheriff,S. and Margolies,M.N. (1993) Modulation of antibody affinity by a non-contact residue. *Protein Sci.*, **2**, 206–214.
Scott,M.G., Crimmins,D.L., McCourt,D.W., Zocher,I., Thiebe,R., Zachau,H.G. and Nahm,M.H. (1989) Clonal characterization of the human IgG antibody repertoire to *Haemophilus influenzae* type b polysaccharide. III. A single V_κII gene and one of several J_κ genes are joined by an invariant arginine to form the most common L chain V region. *J. Immunol.*, **143**, 4110–4116.
Scott,M.G., Crimmins,D.L., McCourt,D.W., Chung,G., Schäble,K.F., Thiebe,R., Quenzel,E.-M., Zachau,H.G. and Nahm,M.N. (1991) Clonal characterization of the human IgG antibody repertoire to *Haemophilus influenzae* type b polysaccharide. IV. The less frequently expressed VL are heterogenous. *J. Immunol.*, **147**, 4007–4013.
Segal,D., Padlan,E., Cohen,G., Rudikoff,S., Potter,M. and Davies,D. (1974) The three dimensional structure of a phosphorylcholine binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Natl Acad. Sci. USA*, **71**, 4298–4302.
Sharon,J. (1990) Structural correlates of high antibody affinity: three engineered amino acid substitutions can increase the affinity of an anti-p-azophenylarsonate antibody 200 fold. *Proc. Natl Acad. Sci. USA*, **87**, 4814–4817.
Sheriff,S., Silverton,E.W., Padlan,E.A., Cohen,G.H., Smith-Gill,S.J., Finzel,B.C. and Davies,D.R. (1987) Three-dimensional structure of an antibody-antigen complex. *Proc. Natl Acad. Sci. USA*, **84**, 8075–8079.
Stanfield,R.L., Takimoto-Kamimura,M., Rini,J.M., Profy,A.T. and Wilson,I.A. (1993) Major antigen induced domain rearrangements in an antibody. *Structure*, **1**, 83–93.
Staubinger,B., Huber,E., Lorenz,W., Osterholzer,E., Pargent,W., Pech,M., Pohlenz,H.-D., Zimmer,F.-J. and Zachau,H.G. (1988a) The human V_κ locus. Characterization of a duplicated region encoding 28 different immunoglobulin genes. *J. Mol. Biol.*, **199**, 23–34.
Staubinger,B., Thiebe,R., Huber,C., Osterholzer,E. and Zachau,H.G. (1988b) Two unusual human immunoglobulin V_κ genes. *Biol. Chem. Hoppe-Seyler*, **369**, 601–607.
Stavnezer,J., Kekish,O., Batter,D., Grenier,J., Balazs,I., Henderson,E. and Zegers,B.J.M. (1985) Aberrant recombination events in B cell lines derived from a κ-deficient human. *Nucleic Acids Res.*, **13**, 3495–3514.
Steipe,B., Plückthun,A. and Huber,R. (1992) Refined crystal structure of a recombinant immunoglobulin domain and a complementarity-determining region 1-grafted mutant. *J. Mol. Biol.*, **225**, 739–753.
Stiernholm,N.B.J. and Berinstein,N.L. (1995) A mutated promoter of a human Ig Vλ gene segment is associated with reduced germ-line transcription and a low frequency of rearrangement. *J. Immunol.*, **154**, 1748–1761.
Suh,S.W., Bhat,T.N., Navia,M.A., Cohen,G.H., Rao,D.N., Rudikoff,S. and Davies,D.R. (1986) The galactan-binding immunoglobulin Fab J539: an x-ray diffraction study at 2.6 Å resolution. *Proteins*, **1**, 74–80.
Tonegawa,S. (1983) Somatic generation of antibody diversity. *Nature*, **302**, 575–581.
Whitlow,M. (1995) Brookhaven Protein Databank (PDB) entry 1FLR.
Wilson,I.A. and Stanfield,R.L. (1993) Antibody-antigen interactions. *Curr. Opin. Struct. Biol.*, **3**, 113–118.
Wilson,I.A. and Stanfield,R.L. (1994) Antibody-antigen interactions: new structures and new conformational changes. *Curr. Opin. Struct. Biol.*, **4**, 857–867.
Zachau,H.G. (1993) The immunoglobulin κ locus—or—what has been learned from looking closely at one-tenth of a percent of the human genome. *Gene*, **135**, 167–173.

Received on April 10, 1995; revised on June 30, 1995

J. Mol. Biol. (1992) 227, 776–798

The Repertoire of Human Germline V_H Sequences Reveals about Fifty Groups of V_H Segments with Different Hypervariable Loops

Ian M. Tomlinson[1,2], Gerald Walter[1,2], James D. Marks[1]
Meirion B. Llewelyn[2] and Greg Winter[1,2]†

[1]*MRC Centre for Protein Engineering and*
[2]*MRC Laboratory of Molecular Biology*
Hills Road, Cambridge CB2 2QH, U.K.

(Received 6 February 1992; accepted 3 June 1992)

We have used the polymerase chain reaction and V_H family-based primers to clone and sequence 74 human germline V_H segments from a single individual and built a directory to include all known germline sequences. The directory contains 122 V_H segments with different nucleotide sequences, 83 of which have open reading frames. The directory indicates that the structural diversity of the germline repertoire for antigen binding is fixed by about 50 groups of V_H segments: each group encodes identical hypervariable loops. The directory should help in mapping the V_H locus, in estimating somatic mutation and V_H segment usage and in designing and constructing synthetic antibody libraries.

Keywords: human antibodies; heavy chain variable region; V_H; polymerase chain reaction

1. Introduction

Antibody architecture accommodates a wealth of structural diversity. Heavy and light chain variable domains (V_H and V_L) each consist of a β-sheet scaffold, surmounted by three antigen-binding loops (complementarity-determining regions, or CDRs‡; Kabut & Wu, 1971) of different lengths which are fleshed with a variety of different side-chains. The structural diversity of the loops can create binding sites of a variety of shapes, ranging from almost flat surfaces (Amit *et al.*, 1986) to deep cavities (Alzari *et al.*, 1990). Underpinning the structural diversity is a combinatorial genetic diversity. For V_H domains, it is generated by the assembly of V_H, D (diversity) and J_H (joining) segments. Two of the CDRs (1 and 2) are encoded by the V_H segment, and CDR3 by the 3′ end of the V_H segment, the D segment and the 5′ end of the J_H segment. With nucleotide addition (N-region diversity at the V_H-D and D-J_H joins), the use of different reading frames in the D segment, and the combination of different rearranged heavy and light chains, the diversity of primary antibody libraries is huge (for reviews, see Tonegawa, 1983; Winter & Milstein, 1991). During an immune response, the antibody variable regions are further diversified by somatic hypermutation, leading to higher affinity binding of the antigen (Berek & Milstein, 1988).

The human V_H, D and J_H segments have been mapped to band q32.33 of chromosome 14 (Croce *et al.*, 1979; Kirsch *et al.*, 1982), and recombine during B cell development. Each V_H segment encodes a 5′ hydrophobic leader peptide and between 95 and 101 amino acid residues of the mature domain flanked at the 3′ end by two recombination signals consisting of a highly conserved heptamer (5′-CACAGTG-3′), a 23-base-pair spacer and a less-conserved nonamer. The V_H segments have evolved by unequal crossing-over, conversion, duplication and deletion (Wysocki & Gefter, 1989; Walter *et al.*, 1990) and can be divided into six families on the basis of nucleotide homology of 80% or above (Kodaira *et al.*, 1986; Lee *et al.*, 1987; Shen *et al.*, 1987; Berman *et al.*, 1988; Humphries *et al.*, 1988; Buluwela & Rabbitts, 1988). The number of V_H segments per individual has most recently been estimated as about 76 (25 V_H1 segments, 5 V_H2 segments, 28 V_H3 segments, 14 V_H4 segments, 3 V_H5 segments and 1 V_H6 segment; Walter *et al.*, 1990), although these figures are likely to be an underestimate (Berman *et al.*, 1988; Walter *et al.*, 1990).

Earlier attempts to clone human V_H segments have involved constructing and probing large cosmid libraries, and have been aimed at mapping and sequencing the whole V_H locus, including pseudogenes (Kodaira *et al.*, 1986; Lee *et al.*, 1987; Berman *et al.*, 1988). In contrast, we set out to

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: CDR, complementarity-determining region; PCR, polymerase chain reaction; FR, framework region; u.v., ultraviolet light.

Table 1
Family-specific primers for PCR amplification of the V_H exon

VH1 primers	
VH1 LEA EX1	5′-CCC AAG CTT CCA TGG ACT GGA CCT GGA G-3′
VH1 LEA EX2	5′-CCC AAG CTT TCA TGG GCT GGA CCT GCA A-3′
VH1 LEA IN	5′-CCC AAG CTT G(A,G)A (A,G)G(A,G) GAT T(G,T) (A,G,T) (G,T)TC CAG T-3′
VH1 LEA EX3	5′-CCC AAG CTT (T,C) (C,T) (C,T) (A,G)CA G(G,A) (T,C,A) (G,A) (C,T) (C,T,G) (C,T)A(C,T,G) (T,G)C-3′
VH1 FR1 (2–8)	5′-CCC AAG CTT (C,G,T)CA(G,A) (C,T)T(A,G,T) (G,T)T(G,A) (C,T)A(G,A) (T,C)C(T,G) G-3′
VH1 FR1 (17–22)	5′-CCC AAG CTT (T,A)C(A,G) G(T,C)G A(A,G) (G,A) (G,A)T(C,T) (T,A)CC TGC-3′
VH1 HEPT	5′-GGA ATT CT(C,G) TGG (G,T)TT (C,T)TC ACA CTG TG-3′
VH2 primers	
VH2 LEA	5′-CCC AAG CTT CTT CTC CAC AGG GGT CTT ATC-3′
VH2 HEPT	5′-GGA ATT CCA CTG TG(C,T) (C,G)CC GCG CAC A-3′
VH3 primers	
VH3 LEA1	5′-CCC AAG CTT T(A,T)(C,T) (A,G)TG TGG CA(A,G,C,T) TTT CTG A-3′
VH3 LEA2	5′-CCC AAG CTT T(A,T) (C,T) (A,G)T(C,G) TG(A,G) (A,C)A(A,G,C,T) TTT CTG A-3′
VH3 LEA3	5′-CCC AAG CTT GT(A,T) TGC A(A,G)G TG(C,T) CCA GTG T-3′
VH3 HEPT	5′-GGA ATT C(A,C)T G(A,G)C (C,T)TC CCC TC(A,G) CT(C,G) TG-3′
VH3 FR1	5′-CCC CCA AGC TTT GT(G,C) CAG (G,C)CT CTG G(A,G)T TC-3′
VH3 FR3	5′-GCT CTA GAG T(G,A)A (G,A)TC (T,G)GC C(T,C)T TCA C(A,G)G-3′
VH3 NON1	5′-GCT CTA GAG GTT TGT G(T,C)C (T,C)GG GC(G,T) CA-3′
VH4 primers	
VH4 LEA	5′-CCC AAG CTT CTG TTC ACA GGG GTC CTG TC-3′
VH4 HEPT	5′-GGA ATT CAC TCA CCT CCC CTC ACT GTG-3′
VH5 primers	
VH5 LEA	5′-CCC AAG CTT AGG TCA CAG AG(A,G) AGA A(C,T)G G-3′
VH5 HEPT	5′-GGA ATT CGC TGG TTT CTC TCA CTG TG-3′
VH6 primers	
VH6 LEA	5′-CCC AAG CTT TCA CAG CAG CAT TCA CAG A-3′
VH6 HEPT	5′-GGA ATT CCT GAC TTC CCC TCA CTG TG-3′

determine the repertoire of human V_H segments that contribute to the structural diversity of the V_H domain. We employed the polymerase chain reaction (PCR) (Saiki *et al.*, 1988) as a method of amplifying individual V_H segments. We designed family-specific primers for V_H segments based on the heptamer and part of the recombination spacer at the 3′ end of the V_H exon, and regions of the leader exon or intron at the 5′ end. Priming from the heptamer has been used to amplify mouse (Borghesi-Nicoletti & Schulze, 1991) and human (Sanz *et al.*, 1989c) V_H segments and has the advantage that since the heptamer is lost during recombination, rearranged V_H genes are not amplified.

2. Materials and Methods

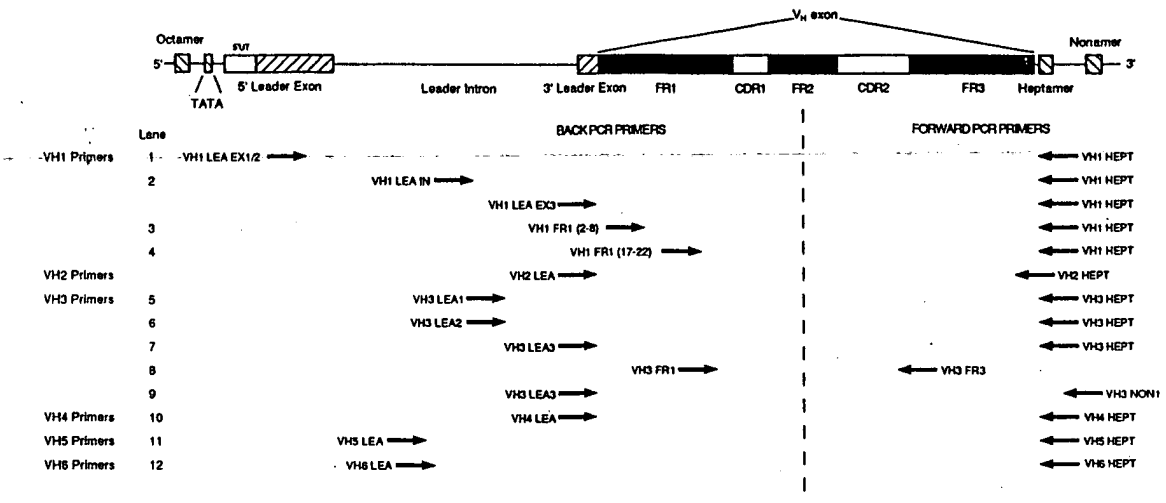
(a) *Primer design*

Primers were designed (Table 1) for each of the 6 V_H families based on the sequences of published V_H segments (Kodaira *et al.*, 1986; Lee *et al.*, 1987; Berman *et al.*, 1988; Humphries *et al.*, 1988) and were located as shown in Fig. 1(a). Forward primers were based around the highly conserved heptamer recombination sequence, 5′-CACAGTG-3′. For 5 V_H families, published germline sequences were used, basing forward primers (VH1 HEPT, VH3 HEPT, VH4 HEPT, VH5 HEPT, VH6 HEPT) on the heptamer sequence and an additional 11 to 13 nucleotides from the recombination spacer. Degenerate nucleotides were incorporated to ensure the efficient priming of known germline genes from each V_H family, and *Eco*RI restriction sites were added for cloning. As germline V_H2 sequences were not available, the forward primer (VH2 HEPT) was designed using the sequence of the third framework (FR) region from a rearranged V_H2 gene, $V_{CE\text{-}1}$ (Takahashi *et al.*, 1984) adding 2 degenerate bases to substitute for those outside FR3, and then adding the conserved heptamer sequence. Family-specific back primers (VH1 LEA EX1, VH1 LEA EX2, VH1 LEA IN, VH1 LEA EX3, VH2 LEA, VH3 LEA1, VH3 LEA2, VH3 LEA3, VH4 LEA, VH5 LEA, VH6 LEA) were based on those parts of the leader exon and intron that are highly conserved within, but not between, V_H families, again incorporating degeneracy where necessary (VH1 LEA EX1 and VH1 LEA EX2 were mixed in equal ratios and are referred to as VH1 LEA EX1/2). The back primers, VH1 FR1 (2–8) and VH1 FR1 (17–22), were subsequently designed using the sequences obtained with the first set of PCR primers. *Hind*III restriction sites were added to all back primers for cloning.

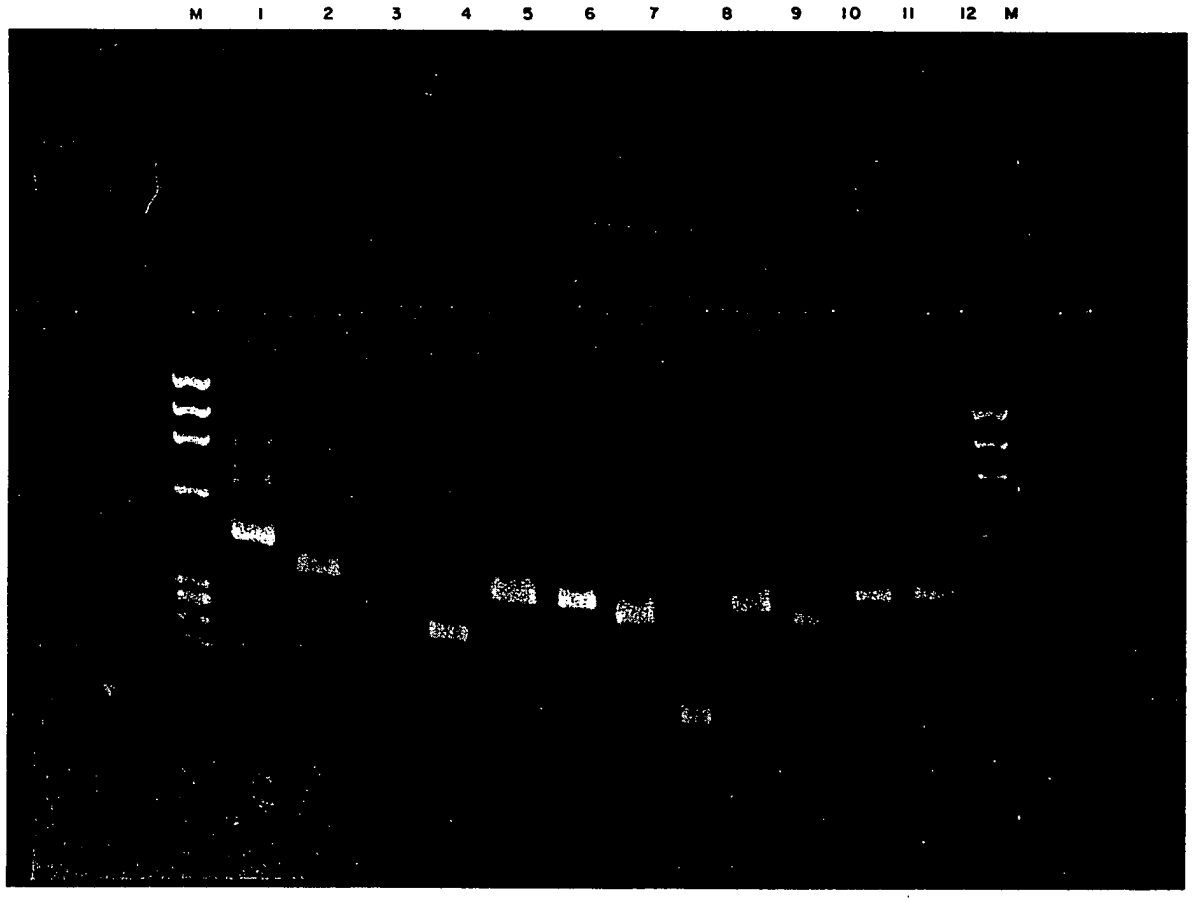
"Internal" primers for the V_H3 family were designed based on those regions of framework 1 (VH3 FR1) and CDR2-framework 3 (VH3 FR3) that display the greatest homology within the V_H3 family (see Fig. 2(b)). Since *Eco*RI restriction sites were noted in 2 published V_H3 pseudogenes ($V_{71\text{-}1}$ and $V_{71\text{-}3}$; Kodaira *et al.*, 1986) we changed the cloning site in the forward primer (VH3 FR3) to *Xba*I.

(b) *Preparation of genomic DNA*

Genomic DNA was isolated from peripheral white blood cells obtained from a healthy Caucasian donor, DP, using a method described by Perry & Carrell (1989). Briefly,



(a)



(b)

Figure 1. Family-specific primers for PCR amplification of the V_H exon. (a) Locations of the family-based PCR primers with respect to the V_H exon. FR, framework region; CDR, complementarity-determining region. Back primers were based in either the leader exon or intron or in framework 1 of the V_H segment. Forward primers were based around the heptamer and nonamer and at the junction of the CDR2 and framework 3. (b) PCR amplified genomic DNA from DP run on a 1·5% agarose gel. M, ϕX174 M_r markers; lanes 1 to 12, amplifications using the sets of primers depicted in (a).

9 ml whole blood was collected in 1 ml 3·8% (w/v) trisodium citrate (anticoagulant). The cells were lysed by adding the mixture to 90 ml ice-cold cell lysis buffer (0·32 M-sucrose, 1% Triton X-100, 5 mM-$MgCl_2$, 10 mM-Tris·HCl (pH 7·5)) and left on ice for 15 min. The nuclear pellet was isolated by centrifugation at 1000 *g* at 4°C for 15 min and then resuspended in 4·5 ml Tris/EDTA (10 mM-Tris·HCl (pH 8·0), 1 mM-EDTA). The pellet was lysed using 10 ml nuclear lysis buffer (0·32 M-lithium acetate, 2% (w/v) SDS, 10 mM-Tris·HCl (pH 8·0), 1 mM-

EDTA), extracted twice with phenol/chloroform, once with chloroform and precipitated using ice-cold ethanol. Samples were resuspended in 500 μl of water and quantified by measuring their absorbance at 260 nm.

(c) *PCR amplification and sequencing*

Primers were synthesized on an Applied Biosystems (Foster City, CA) oligonucleotide synthesizer. Genomic DNA was amplified using the pairs of PCR primers (Fig. 1(a) and Table 1) in a Techne programmable Dri-Block PHC-1 thermal cycler (Cambridge, UK) with either Promega (Madison, WI) or Cetus (Perkin Elmer, Norwalk, CT) *Thermus aquaticus* (Taq) DNA polymerase. Reaction mixtures (50 μl) were prepared containing 25 pmol of each primer, 5 to 10 μg of genomic DNA, 2·5 units of Taq polymerase, 200 μM (each) dNTPs and the recommended buffer (Promega: 50 mM-KCl, 10 mM-Tris·HCl (pH 8·8), 1·5 mM-$MgCl_2$, 0·1% Triton X-100; Cetus: 50 mM-KCl, 10 mM-Tris·HCl (pH 8·3), 1·5 mM-$MgCl_2$, 0·001% (w/v) gelatin). The reaction mixture was overlaid with paraffin oil and 30 cycles of amplification were performed. Each cycle consisted of denaturation (94°C for 1 min), annealing (55°C for 1 min) and extension (72°C for 2 min). At the end of 30 cycles, there was a final extension at 65°C for 5 min. The product was analysed by running 5 μl on a 1·5% (w/v) agarose gel. The remainder was extracted with phenol/chloroform, precipitated with ethanol and digested with restriction enzymes *Hin*dIII and *Eco*RI (or *Xba*I). A band of the expected size was cut from a 1·5% low melting point agarose gel and then purified by adsorption onto glassmilk using Geneclean II (Bio 101, La Jolla, CA) or by electroelution followed by precipitation with ethanol.

The product was ligated into M13-K19 (Carter *et al.*, 1985) that had been digested with *Hin*dIII and *Eco*RI (or *Xba*I). The ligation mix was used to transform *E. coli* BMH 71-18 cells (Gronenborn, 1976) by electroporation (Dower *et al.*, 1988) using the Bio-Rad (Richmond, CA) Gene Pulser and plated on TYE plates (Miller, 1972). Single-stranded template from selected plaques was prepared and sequenced using the dideoxy chain termination method (Sanger *et al.*, 1977) and modified T7 DNA polymerase (Sequenase II, United States Biochemical Corp., Cleveland, Ohio). The sequence was read in one direction and compressions resolved using deoxyinosine triphosphate (Mills & Kramer, 1979).

Several precautions were taken to avoid cross-contamination. PCR reaction mixes were subjected to high intensity short-wave u.v. radiation (Amplirad, Genetic Research Instrumentation, Dunmow, Essex, U.K.) for 5 min before adding genomic DNA to destroy any DNA contamination. Negative controls (no genomic DNA added) were always included in all amplifications to check for DNA contamination. Independent amplifications with identical sets of primers were undertaken simultaneously to avoid clones isolated from one amplification contaminating the next. In all cases we imposed the requirement that each germline V_H segment was seen in at least 2 independent amplifications.

(d) *Probing*

Oligonucleotide probes, 17 to 21 nucleotides in length (Table 2) were designed as described in Results, and synthesized as above. Phage plaques were picked onto duplicate TYE plates and grown as colonies for 30 h at 37°C. (Plaques that should hybridize to the probes were always included as positive controls.) The colonies were lifted onto Hybond nylon filters (Amersham Int., Amersham, U.K.), denatured in 5% (w/v) SDS. 2× SSC (300 mM-NaCl, 30 mM-trisodium citrate, pH 7·0) for 2 min, baked in a microwave oven for 2·5 min and autocrosslinked by short-wave u.v. (Stratalinker: Stratagene, La Jolla, CA) (Buluwela *et al.*, 1989). Filters were prehybridized for 20 min at 42°C in 15 ml hybridization solution (1 M-NaCl, 1× Denhardt's (0·02% Ficoll, 0·02% polyvinylpyrrolidone, 0·02% bovine serum albumin), 100 mM-Tris·HCl (pH 7·5), 6·25 mM-EDTA, 1 mM-sodium pyrophosphate, 0·5% Nonidet P40, 0·006% rATP, 0·02% brewers' yeast tRNA) using a Techne HB-1 Hybridiser (Cambridge, U.K.).

For probing, 15 pmol of oligonucleotide were phosphorylated with 30 μCi [^{32}P]dATP for 30 min using 2 units of polynucleotide kinase (New England Biolabs, Beverly, MA) in 30 μl 50 mM-Tris·HCl (pH 7·5), 10 mM-$MgCl_2$, 1 mM-dithiothreitol, and incorporation of ^{32}P checked by electrophoresis of the oligonucleotide on an 18% (w/v) polyacrylamide gel. The probe was added to the hybridization solution, and the filters were hybridized at 42°C for 2 h and then washed with 40 ml 6× SSC (see above), 0·1% SDS, 0·1% sodium pyrophosphate at this temperature for 15 min and then for 20 min with 40 ml 3 M-TMACl (tetramethylammonium chloride) in 50 mM-Tris·HCl (pH 8·0), 0·1% SDS and 2 mM-EDTA (Wood *et al.*, 1985) at 59°C (17-mer), 61°C (18-mer), 63°C (19-mer) or 67°C (21-mer). Filters were dried and exposed to Kodak Fast Film overnight using an intensifying screen at −70°C. Filters were recycled by washing at 80 to 90°C for 5 min in 2× SSC and could be probed several times without loss of signal.

(e) *Compilation of germline and rearranged V_H database*

DNA sequences were aligned and translated by a sequence analysis program (MacVector, IBI Kodak, New Haven, CT). In order to compile a comprehensive database of both human germline and rearranged V_H sequences we searched MedLine (U.S. National Library of Medicine), GenBank (IntelliGenetics Inc., Real Mountain View, CA) and Kabat (Proteins of Immunological Interest, Kabat *et al.*, 1991) databases (for references, see Figs 2 and 3) and incorporated our own data. Rearranged genes were assigned to their closest germline counterparts by the presence of specific motifs in the protein sequence indicative of a particular V_H segment or by maximum homology of the nucleotide sequences (using MacVector).

3. Results

(a) *Strategy*

We designed family-specific PCR primers based on sequences from the literature and amplified, cloned and sequenced germline V_H segments from our donor DP. Nucleotide sequences were aligned and taken as confirmed when seen as identical in two independent amplifications. Genes which remained unconfirmed in phase 1 were probed for with ^{32}P-labelled oligonucleotides and sequenced in phase 2.

(b) *Phase 1: PCR amplification and sequencing of random clones*

Genomic DNA was amplified using sets of family-based primers. The majority of primer combinations

Table 2

Oligonucleotide probes used for identification of germline V_H segments

V_H1 family	
DP-1	5′-AGT AAT ACG TGG CCG TG-3′
DP-1/8	5′-TGT GCC ACC ACT GTT AG-3′
DP-2	5′-CCA CTG CCA ACG ACG AT-3′
DP-3/5/24	5′-ATT GTT TCA CCA TCT TC-3′
DP-4	5′-TGC AGG TAG CGG TAG GT-3′
DP-4	5′-ACC ATT GAA AGG TGT GA-3′
DP-5	5′-TGG ATA ATT CAG TGA GG-3′
DP-6	5′-ACT GTG TAA AGT ATT TG-3′
DP-7/22	5′-CAG TGC ATA TAG TAG CT-3′
DP-8	5′-TCG TCA GAT CTC AGC CT-3′
DP-9/10	5′-CCA GCT GAT AGC ATA GC-3′
DP-9/21	5′-GGT TCC CAG TGT TGG TG-3′
DP-10	5′-TGC TGT ACC AAA GAT AG-3′
DP-11	5′-AGG TGT ATC CAC AAG TCT-3′
DP-12	5′-ATC ACT AGG GCA CAC CAA-3′
DP-13	5′-ACA TTG GGT TCA CCA GGG-3′
DP-14/22	5′-TGT GTT ACC ATT GTA AG-3′
DP-15	5′-AGT TGA TAT CAT AAC TG-3′
DP-16/17/20	5′-TTG CCA GAG TAG CTC CC-3′
DP-18	5′-GAT CTG AAG ACA CGC CG-3′
DP-19	5′-GAC TAC ACC AGT TGG AC-3′
DP-19	5′-GTT CAT AAA GTA GTC GG-3′
DP-19	5′-TGC TCG AAG ATG TGT CC-3′
DP-19/23/25	5′-GTG TTA CCA TTG CCA GC-3′
DP-21	5′-CAA CTC AGA CCC AGA TT-3′
DP-22	5′-CGG CCA TGT CGT CAG AT-3′
DP-23	5′-GCA TAA AGT TGT TGG TG-3′
DP-24	5′-CCC AGG TTT CCT CAC CT-3′
DP-1/7/8/10/14/19/21/22/23/25	5′-CAC TGT GTC TCT CGC AC-3′
Rearranged gene probes	
333. 1H1, etc.	5′-CAG TGT ATG GTG GAG TCA C-3′
VDJ191	5′-AGT CAG GGC ATG ATT ATT A-3′
39-1	5′-GCC CAC ACC CAC TCC ACT AGT-3′
41-1	5′-GCC CAC ACC CCC TCC ACT AGT-3′
V_H3 family	
DP-29	5′-TTG TTT CTA GTA CGG CCA A-3′
DP-30	5′-TTC TTA TTA AAC CTA CCA A-3′
DP-31	5′-CAC TAT TCC AAC TAA TAC C-3′
DP-32	5′-GTG CTA CCA CCA TTC CAA T-3′
DP-33	5′-CAC CAT CCC AAC TAA TAA G-3′
DP-36	5′-AGC TTT GCT TTT AAT ACA G-3′
DP-37	5′-AGC TTT GCT TTT AAT ACG G-3′
DP-41	5′-GTG CAT GCC ATA GTT ACT G-3′
DP-42	5′-TAC CAC CGC TAT AAA TAA C-3′
DP-44/45	5′-GTG CCA CCA CCA GTA CCA A-3′
DP-44/45/46/61	5′-CAG TGC ATA GCA TAG CTA C-3′
DP-47	5′-CCA CTA CCA CTA ATA GCT G-3′
DP-49/50	5′-CAG TGC ATG CCA TAG CTA C-3′
DP-46/49	5′-TCA TAT GAT ATA ACT GCC A-3′
DP-50	5′-TCA TAC CAT ATA ACT GCC A-3′
DP-51	5′-CAG TTC ATG CTA TAG CTA C-3′
DP-52	5′-CAG TGC AGA ACA TAG CTA C-3′
DP-53	5′-CCA TCA CTA TTA ATA CGT G-3′
DP-54	5′-TTC CAT CTT GCT TTA TGT T-3′
DP-55/56	5′-CCC CAT TAG GAT TAA CTT G-3′
DP-58	5′-AGT TCA TTT CAT AAC TA-3′
DP-59	5′-CAG TTC ATG TCA CTG TTA C-3′
DP-60	5′-GTA GCC ATA GCA CGC ACT G-3′
DP-61	5′-ACC CCC ATT ACT ACT AAT A-3′
V_H4 family	
$V_{2.1}$	5′-GCC CCA GTA GTA ACT ACT ACT-3′
V_{58}	5′-GTA GTA ACC ACT GAC GGA C-3′
V_{11}	5′-AGT TGG GGT TCC CAC TAT G-3′
V_{79}	5′-GGT CCC CGG AGG CTT CAC C-3′

produced good intensity PCR bands, as is shown in Figure 1(b), but amplifications using VH1 EX3/VH1 HEPT and VH2 LEA/VH2 HEPT were variable and hence are not shown. Initially, 596 random clones were sequenced (V_H1 family (170), V_H2 family (120), V_H3 family (150), V_H4 family (120), V_H5 family (24) and V_H6 family (12)). With one exception (one V_H5 gene found in a V_H1 library), the primers proved family-specific. This initial round of sequencing established 35 V_H sequences (including pseudogenes) that were identical in at least *two independent* PCR amplifications (V_H1 family (12), V_H2 family (3), V_H3 family (8), V_H4 family (10), V_H5 family (1) and V_H6 family (1)) and by this criterion correspond to germline V_H segments.

Many sequences were unconfirmed due to single nucleotide differences between clones from independent amplifications, presumably due to errors introduced by the Taq polymerase. The 61 single base changes seen per 100 sequences for the V_H1 and V_H3 families correspond to 7×10^{-5} changes/nucleotide per cycle, which is consistent with the Taq polymerase error rate suggested by Maruyama (1990).

Other sequences, never confirmed in independent amplifications (but sometimes found in more than one clone from the same amplification), consisted of two parts, each of which could be aligned to different V_H segments. As became clear on probing (see below), these sequences arose from partially extended fragments reannealing to a different segment after heat-denaturation. This phenomenon, termed "PCR cross-over", has also been seen in the detection of homologous recombinants (Frohman & Martin, 1990) and in the amplification of preproinsulin cDNA (Shuldiner *et al.*, 1989) and in this study accounted for 10% of all V_H1 and V_H3 clones sequenced.

For the smaller V_H families (V_H2, V_H4, V_H5, V_H6), all sequences were confirmed in phase 1, or could be explained by PCR artifacts. But many sequences from the V_H1 and V_H3 families remained unconfirmed, requiring systematic probing of a larger number of clones.

(c) *Phase 2: probing and directed sequencing*

With the V_H1 primers, 42 different sequences (excluding obvious PCR errors caused by single base substitutions) were obtained in phase 1. Only 12 of these sequences were identical in at least two independent amplifications. Therefore, motif-specific probes were designed (Table 2) such that each probe would identify a group of different V_H1 clones with a particular sequence motif. Hence, when each clone

was probed in turn with each of the 29 probes, it could be distinguished by its "fingerprint", i.e. the set of sequence motifs that it contains. Thus, 1750 clones from independent amplifications using the five V_H1-based primer combinations (Fig. 1) were regridded and hybridized with the 29 probes. Clones that appeared to confirm a sequence from phase 1 by "fingerprinting" were sequenced. In this way a further 11 V_H1 sequences were confirmed and only two new (pseudo)genes (DP-17, DP-20) were discovered. Nineteen of the original 42 sequences could not be confirmed by probing, but 18 of these could be attributed to "PCR cross-over".

For the majority of unconfirmed sequences in the V_H3 family, we designed gene-specific probes (17- and 19-mers, Table 2), except in the case of DP-46/DP-49, where three probes were necessary for identification, and DP-44/45 and DP-56/57, where discrimination between the two in each pair was not possible. Probes were centred on the region of greatest heterogeneity within a CDR and therefore a single probe (with the above exceptions) could identify a single V_H segment. Thus, 1100 clones taken from independent amplifications with the three sets of V_H3 leader/heptamer-based primers (Fig. 1) were hybridized in turn with the 21 probes and a further 22 V_H3 segments were confirmed by directed sequencing. The remaining unconfirmed sequences could be attributed to PCR artifacts.

We also designed "internal" VH3 primers (VH3 FR1 and VH3 FR3) based on sequence data from phase 1 and phase 2. Genomic DNA from DP was amplified as before, and 48 randomly selected clones were sequenced and confirmed, when necessary, in two independent amplifications by probing and directed sequencing. Only seven new V_H segments were obtained, three of which appeared to be fragmented pseudogenes with less than 60% homology to any known V_H segment. Two sequences had been published before and have unusual heptamer sequences (DP-59/V_H19 and DP-62/V_{71-1}, respectively) and the other two sequences were new (DP-60 and DP-61).

To isolate full length versions of genes DP-59 to DP-61, which have open reading frames, we designed a primer (VH3 NON1) based on nonamer sequences of V_H segments reviewed by Pascual & Capra (1991). Amplifications of genomic DNA were performed using VH3 LEA3 and VH3 NON1, and the resulting fragments were cloned, regridded and probed with oligonucleotides specific for DP-59, DP-60 and DP-61. DP-59 and DP-60 were isolated from independent PCR amplifications, and shown to have unusual heptamer sequences. A full length version of DP-61 was not found in this library.

We also attempted to confirm additional germline V_H segments reported in the literature and germline analogues of published rearranged genes. Using the V_H family-specific primers (Table 1) to amplify and clone germline V_H2, V_H3 and V_H4 segments, we probed (Table 2) for the germline V_H segments V_{11}, V_{58}, V_{79} and V_{2-1} (Lee *et al.*, 1987), rearranged V_H genes 39-1, 41-1 (Deane & Norton, 1990), VDJ191 (Mensink *et al.*, 1986) and 333, 1H1, 2C12, 2A12, 1B11, 112, 115 and 126 (Cleary *et al.*, 1986) (rearranged genes were probed for at low stringency, i.e. TMACl wash at 50°C). None of these genes was identified in our libraries.

(d) *Sequence directory*

The 74 germline V_H segments (25 V_H1 segments, 3 V_H2 segments, 34 V_H3 segments, 10 V_H4 segments, 1 V_H5 segment and 1 V_H6 segment) cloned and sequenced by us are prefixed "DP", the initials of our donor and are denoted by running numbers. Of these, 51 have open reading frames and 23 contain either frame shifts or stop codons and are therefore considered to be pseudogenes. We have also included sequences of germline V_H segments published by others. The protein and nucleotide sequences of all 83 germline V_H segments with open reading frames are given in Figure 2(a) and (b), respectively, and nucleotide sequences of the 39 germline V_H segments with interrupted reading frames (either frame shifts or stop codons) in Figure 2(c). In Figure 2(b), the nucleotide sequences in each family have been aligned to a sequence with an open reading frame, 21-2 (V_H1 family), V_{II-5} (V_H2 family), 12-2 (V_H3 family), V_{71-2} (V_H4 family), VH251 (V_H5 family), V_H-VI (V_H6 family). The same sequences were used to align the pseudogenes in Figure 2(c).

"f1-p1" is a V_H segment described by Olee *et al.* (1991), which was seen in amplifications of genomic DNA from two individuals, Fer and Pla. The V_H segments hv3005b54, hv3019b13, hv3019b18 (Olee *et al.*, 1991), $V_H4.12$, $V_H4.14$, $V_H4.15$ (Sanz *et al.*, 1989c) are genes amplified by PCR, but not confirmed either by probing, independent amplifications, a rearranged sequence or by independent work. These sequences may be the result of PCR artifacts and have therefore been excluded from Figure 2.

Within each family, protein sequences are arranged alphabetically by the amino acid residues (single letter code) of CDR1 and where these are identical by CDR2 (Fig. 2(a)). Sequences with minor framework differences, which could include allelic differences, are therefore adjacent. Sequences with identical encoded CDRs 1 and 2 are grouped with brackets (these also have identical H1 and H2 hypervariable loops, as defined by Chothia *et al.* (1992), except in the case of 21-2/3-1/DP-7 and HG3; and $V_H4.11$/DP-71, V_{71-4} and $V_H4.16$). The canonical structure classes of H1 (CDR1) and H2 (CDR2) (Chothia & Lesk, 1987; Chothia *et al.*, 1989, 1992) are shown, and those sequences that may be defective on structural grounds are marked with an *X* (see Chothia *et al.*, 1992). The canonical structure class of DP-61 is unknown.

V_H segments that have heptamers other than the conserved 5'-CACAGTG-3' motif are marked *H*. The nonamer is generally conserved within each

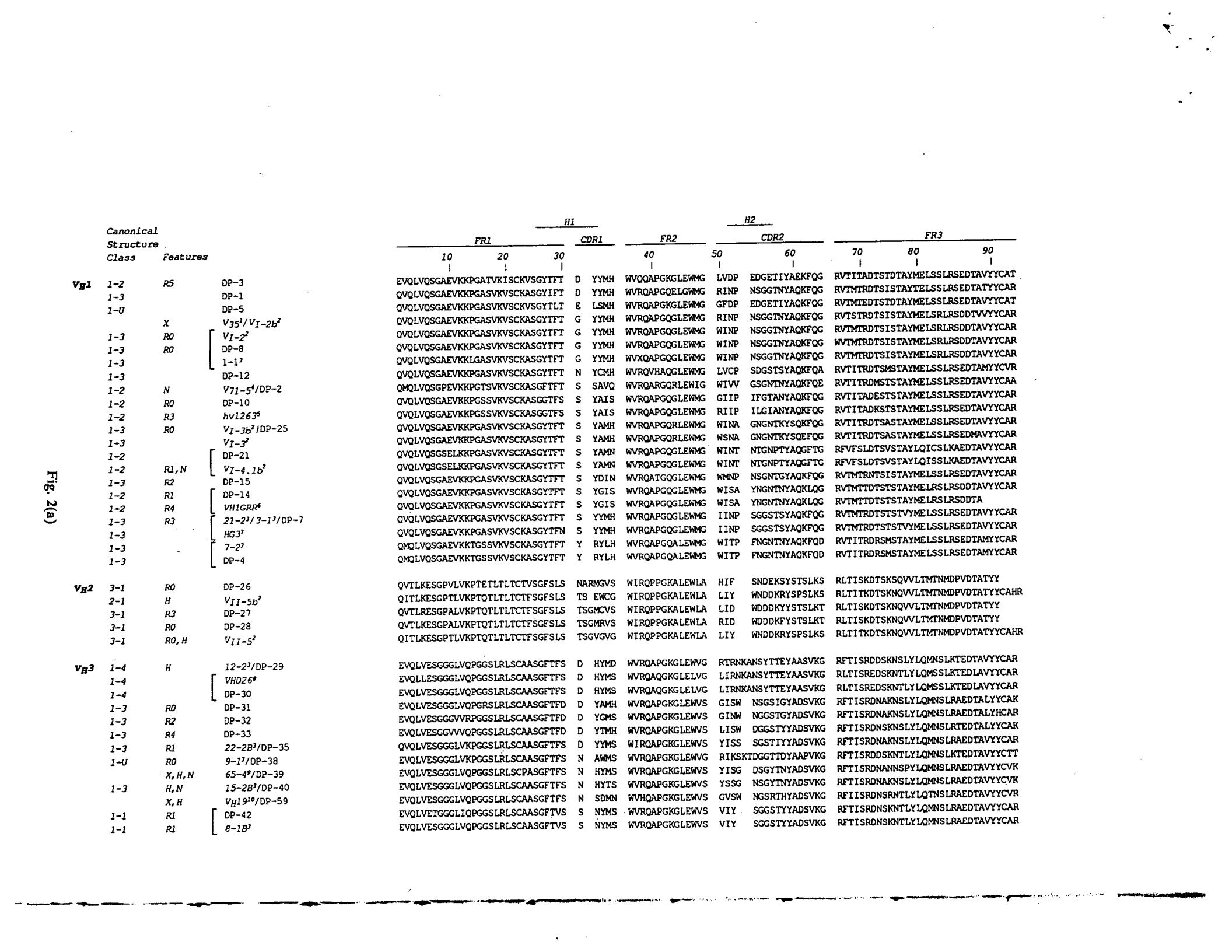

Canonical Structure Class
Features
H1
H2
FR1
CDR1
FR2
CDR2
FR3
10
20
30
40
50
60
70
80
90
VH1
1-2
R5
DP-3
EVQLVQSGAEVKKPGATVKISCKVSGYTFT D YYMH WVQQAPGKGLEWMG LVDP EDGETIYAEKFQG RVTITADTSTDTAYMELSSLRSEDTAVYYCAT
1-3
DP-1
QVQLVQSGAEVKKPGASVKVSCKASGYIFT D YYMH WVRQAPGQELGWMG RINP NSGGTNYAQKFQG RVTMTRDTSISTAYTELSSLRSEDTATYYCAR
1-U
DP-5
QVQLVQSGAEVKKPGASVKVSCKVSGYTLT E LSMH WVRQAPGKGLEWMG GFDP EDGETIYAQKFQG RVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT
X
V35[1]/VI-2b[2]
QVQLVQSGAEVKKPGASVKVSCKASGYTFT G YYMH WVRQAPGQGLEWMG RINP NSGGTNYAQKFQG RVTSTRDTSISTAYMELSRLRSDDTVVYYCAR
1-3
R0
VI-2[2]
QVQLVQSGAEVKKPGASVKVSCKASGYTFT G YYMH WVRQAPGQGLEWMG WINP NSGGTNYAQKFQG RVTMTRDTSISTAYMELSRLRSDDTAVYYCAR
1-3
R0
DP-8
QVQLVQSGAEVKKPGASVKVSCKASGYTFT G YYMH WVRQAPGQGLEWMG WINP NSGGTNYAQKFQG WVTMTRDTSISTAYMELSRLRSDDTAVYYCAR
1-3
1-1[3]
QVQLVQSGAEVKKLGASVKVSCKASGYTFT G YYMH WVXQAPGQGLEWMG WINP NSGGTNYAQKFQG RVTMTRDTSISTAYMELSRLRSDDTAVYYCAR
1-3
DP-12
QVQLVQSGAEVKKPGASVKVSCKASGYTFT N YCMH WVRQVHAQGLEWMG LVCP SDGSTSYAQKFQA RVTITRDTSMSTAYMELSSLRSEDTAMYYCVR
1-2
N
V71-5[4]/DP-2
QMQLVQSGPEVKKPGTSVKVSCKASGFTFT S SAVQ WVRQARGQRLEWIG WIVV GSGNTNYAQKFQE RVTITRDMSTSTAYMELSSLRSEDTAVYYCAA
1-2
R0
DP-10
QVQLVQSGAEVKKPGSSVKVSCKASGGTFS S YAIS WVRQAPGQGLEWMG GIIP IFGTANYAQKFQG RVTITADESTSTAYMELSSLRSEDTAVYYCAR
1-2
R3
hv1263[5]
QVQLVQSGAEVKKPGSSVKVSCKASGGTFS S YAIS WVRQAPGQGLEWMG RIIP ILGIANYAQKFQG RVTITADKSTSTAYMELSSLRSEDTAVYYCAR
1-3
R0
VI-3b[2]/DP-25
QVQLVQSGAEVKKPGASVKVSCKASGYTFT S YAMH WVRQAPGQRLEWMG WINA GNGNTKYSQKFQG RVTITRDTSASTAYMELSSLRSEDTAVYYCAR
1-3
VI-3[2]
QVQLVQSGAEVKKPGASVKVSCKASGYTFT S YAMH WVRQAPGQRLEWMG WSNA GNGNTKYSQEFQG RVTITRDTSASTAYMELSSLRSEDMAVYYCAR
1-2
DP-21
QVQLVQSGSELKKPGASVKVSCKASGYTFT S YAMN WVRQAPGQGLEWMG WINT NTGNPTYAQGFTG RFVFSLDTSVSTAYLQICSLKAEDTAVYYCAR
1-2
R1,N
VI-4.1b[2]
QVQLVQSGSELKKPGASVKVSCKASGYTFT S YAMN WVRQAPGQGLEWMG WINT NTGNPTYAQGFTG RFVFSLDTSVSTAYLQISSLKAEDTAVYYCAR
1-3
R2
DP-15
QVQLVQSGAEVKKPGASVKVSCKASGYTFT S YDIN WVRQATGQGLEWMG WMNP NSGNTGYAQKFQG RVTMTRNTSISTAYMELSSLRSEDTAVYYCAR
1-2
R1
DP-14
QVQLVQSGAEVKKPGASVKVSCKASGYTFT S YGIS WVRQAPGQGLEWMG WISA YNGNTNYAQKLQG RVTMTTDTSTSTAYMELRSLRSDDTAVYYCAR
1-2
R4
VH1GRR[6]
QVQLVQSGAEVKKPGASVKVSCKASGYTFT S YGIS WVRQAPGQGLEWMG WISA YNGNTNYAQKLQG RVTMTTDTSTSTAYMELRSLRSDDTA
1-3
R3
21-2[3]/3-1[3]/DP-7
QVQLVQSGAEVKKPGASVKVSCKASGYTFT S YYMH WVRQAPGQGLEWMG IINP SGGSTSYAQKFQG RVTMTRDTSTSTVYMELSSLRSEDTAVYYCAR
1-3
HG3[7]
QVQLVQSGAEVKKPGASVKVSCKASGYTFN S YYMH WVRQAPGQGLEWMG IINP SGGSTSYAQKFQG RVTMTRDTSTSTVYMELSSLRSEDTAVYYCAR
1-3
7-2[3]
QMQLVQSGAEVKKTGSSVKVSCKASGYTFT Y RYLH WVRQAPGQALEWMG WITP FNGNTNYAQKFQD RVTITRDRSMSTAYMELSSLRSEDTAMYYCAR
1-3
DP-4
QMQLVQSGAEVKKTGSSVKVSCKASGYTFT Y RYLH WVRQAPGQALEWMG WITP FNGNTNYAQKFQD RVTITRDRSMSTAYMELSSLRSEDTAMYYCAR
VH2
3-1
R0
DP-26
QVTLKESGPVLVKPTETLTLTCTVSGFSLS NARMGVS WIRQPPGKALEWLA HIF SNDEKSYSTSLKS RLTISKDTSKSQVVLTMTNMDPVDTATYY
2-1
H
VII-5b[2]
QITLKESGPTLVKPTQTLTLTCTFSGFSLS TS EWCG WIRQPPGKALEWLA LIY WNDDKRYSPSLKS RLTITKDTSKNQVVLTMTNMDPVDTATYYCAHR
3-1
R3
DP-27
QVTLRESGPALVKPTQTLTLTCTFSGFSLS TSGMCVS WIRQPPGKALEWLA LID WDDDKYYSTSLKT RLTISKDTSKNQVVLTMTNMDPVDTATYY
3-1
R0
DP-28
QVTLKESGPALVKPTQTLTLTCTFSGFSLS TSGMRVS WIRQPPGKALEWLA RID WDDDKFYSTSLKT RLTISKDTSKNQVVLTMTNMDPVDTATYY
3-1
R0,H
VII-5[2]
QITLKESGPTLVKPTQTLTLTCTFSGFSLS TSGVGVG WIRQPPGKALEWLA LIY WNDDKRYSPSLKS RLTITKDTSKNQVVLTMTNMDPVDTATYYCAHR
VH3
1-4
H
12-2[3]/DP-29
EVQLVESGGGLVQPGGSLRLSCAASGFTFS D HYMD WVRQAPGKGLEWVG RTRNKANSYTTEYAASVKG RFTISRDDSKNSLYLQMNSLKTEDTAVYYCAR
1-4
VHD26[8]
EVQLLESGGGLVQPGGSLRLSCAASGFTFS D HYMS WVRQAQGKGLELVG LIRNKANSYTTEYAASVKG RLTISREDSKNTLYLQMSSLKTEDLAVYYCAR
1-4
DP-30
EVQLVESGGGLVQPGGSLRLSCAASGFTFS D HYMS WVRQAQGKGLELVG LIRNKANSYTTEYAASVKG RLTISREDSKNTLYLQMSSLKTEDLAVYYCAR
1-3
R0
DP-31
EVQLVESGGGLVQPGRSLRLSCAASGFTFD D YAMH WVRQAPGKGLEWVS GISW NSGSIGYADSVKG RFTISRDNAKNSLYLQMNSLRAEDTALYYCAK
1-3
R2
DP-32
EVQLVESGGGVVRPGGSLRLSCAASGFTFD D YGMS WVRQAPGKGLEWVS GINW NGGSTGYADSVKG RFTISRDNAKNSLYLQMNSLRAEDTALYHCAR
1-3
R4
DP-33
EVQLVESGGVVVQPGGSLRLSCAASGFTFD D YTMH WVRQAPGKGLEWVS LISW DGGSTYYADSVKG RFTISRDNSKNSLYLQMNSLRTEDTALYYCAK
1-3
R1
22-2B[3]/DP-35
QVQLVESGGGLVKPGGSLRLSCAASGFTFS D YYMS WIRQAPGKGLEWVS YISS SGSTIYYADSVKG RFTISRDNAKNSLYLQMNSLRAEDTAVYYCAR
1-U
R0
9-1[3]/DP-38
EVQLVESGGGLVKPGGSLRLSCAASGFTFS N AWMS WVRQAPGKGLEWVG RIKSKTDGGTTDYAAPVKG RFTISRDDSKNTLYLQMNSLKTEDTAVYYCTT
X,H,N
65-4[9]/DP-39
EVQLVESGGGLVQPGGSLRLSCPASGFTFS N HYMS WVRQAPGKGLEWVS YISG DSGYTNYADSVKG RFTISRDNANNSPYLQMNSLRAEDTAVYYCVK
1-3
H,N
15-2B[3]/DP-40
EVQLVESGGGLVQPGGSLRLSCAASGFTFS N HYTS WVRQAPGKGLEWVS YSSG NSGYTNYADSVKG RFTISRDNAKNSLYLQMNSLRAEDTAVYYCVK
X,H
VH19[10]/DP-59
EVQLVESGGGLVQPGGSLRLSCAASGFTFS N SDMN WVHQAPGKGLEWVS GVSW NGSRTHYADSVKG RFIISRDNSRNTLYLQTNSLRAEDTAVYYCVR
1-1
R1
DP-42
EVQLVETGGGLIQPGGSLRLSCAASGFTVS S NYMS WVRQAPGKGLEWVS VIY SGGSTYYADSVKG RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAR
1-1
R1
8-1B[3]
EVQLVESGGGLVQPGGSLRLSCAASGFTVS S NYMS WVRQAPGKGLEWVS VIY SGGSTYYADSVKG RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAR

Fig. 2(a)

		X	65-2[9]/DP-44	EVQLVHSGGGLVQPGGSLRLSCAGSGFTFS	S	YAMH	WVRQAPGKGLEWVS	AIG	TGGGTYYADSVKG	RFTISRDNAKNSLYLQMNSLRAEDMAVYYCAR
	1-1		DP-45	EVQLVQSGGGLVQPGGSLRLSCAGSGFTFS	S	YAMH	WVRQAPGKGLEWVS	AIG	TGGGTYYADSVKG	RFTISRDNAKNSLYLQMNSLRAEDMAVYYCAR
	1-3	R10	f1-p1[11]	EVQLVESGGGLVQPGGSLRLSCSASGFTFS	S	YAMH	WVRQAPGKGLEYVS	AISS	NGGSTYYADSVKG	RFTISRDNSKNTLYVQMSSLRAEDTAVYYCVR
	.		DP-61	TFS	S	YAMH	WVRQAPGKGLEYVS	AISS	NGGSTYYAD	
	1-3		hv3005[12]	QVQLVESGGGVVQPGRSLRLSCAASGFTFS	S	YAMH	WVRQAPGKGLEWVA	VISY	DGSNKYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAR
	1-3	R0	hv3005f3[11]	QVQLVESGGGVVQPGRSLRLSCAASGFTFS	S	YAMH	WVRQAPGKGLEWVA	VISY	DGSNKYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAR
	1-3	R0	GL-SJ2[13]/DP-46	QVQLVESGGGVVQPGRSLRLSCAASGFTFS	S	YAMH	WVRQAPGKGLEWVA	VISY	DGSNKYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAR
	1-3	R0	VH26[14]/DP-47	EVQLLESGGGLVQPGGSLRLSCAASGFTFS	S	YAMS	WVRQAPGKGLEWVS	AISG	SGGSTYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK
	1-1	R0	13-2[3]/DP-48	EVQLVESGGGLVQPGGSLRLSCAASGFTFS	S	YDMH	WVRQATGKGLEWVS	AIG	TAGDTYYPGSVKG	RFTISRENAKNSLYLQMNSLRAGDTAVYYCAR
	1-3	R0	DP-58	EVQLVESGGGLVQPGGSLRLSCAASGFTFS	S	YEMN	WVRQAPGKGLEWVS	YISS	SGSTIYYADSVKG	RFTISRDNAKNSLYLQMNSLRAEDTAVYYCAR
	1-3	R0	1.9III[3]/DP-49	QVQLVESGGGVVQPGRSLRLSCAASGFTFS	S	YGMH	WVRQAPGKGLEWVA	VISY	DGSNKYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK
	1-3	R0	3019b9[11]/DP-50	QVQLVESGGGVVQPGRSLRLSCAASGFTFS	S	YGMH	WVRQAPGKGLEWVA	VIWY	DGSNKYYADSVKG	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAR
	1-3	R3	DP-51	EVQLVESGGGLVQPGGSLRLSCAASGFTFS	S	YSMN	WVRQAPGKGLEWVS	YISS	SSSTIYYADSVKG	RFTISRDNAKNSLYLQMNSLRDEDTAVYYCAR
		X	DP-52	EDQLVESGGGLVQPGGSLRPSCAASGFAFS	S	YVLH	WVRRAPGKGPEWVS	AIG	TGGDTYYADSVMG	RFTISRDNAKKSLYLQMNSLIAEDMAVYYCAR
	1-3	R2	H11[15]/DP-53	EVQLVESGGGLVQPGGSLRLSCAASGFTFS	S	YWMH	WVRQAPGKGLVWVS	RINS	DGSSTTYADSVKG	RFTISRDNAKNTLYLQMNSLRAEDTAVYYCAR
	1-3	R0	DP-54	EVQLVESGGGLVQPGGSLRLSCAASGFTFS	S	YWMS	WVRQAPGKGLEWVA	NIKQ	DGSEKYYVDSVKG	RFTISRDNAKNSLYLQMNSLRAEDTAVYYCAR
V_H4	1-1	R2	Tou-VH4.21[16]	QVQLQQWGAGLLKPSETLSLTCAVYGGSFS	G	YYWS	WIRQPPGKGLEWIG	EII	HSGSTNYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	1-1	R0	$V_H5^{10}/V_H4.21^{17}$/DP-63	QVQLQQWGAGLLKPSETLSLTCAVYGGSFS	G	YYWS	WIRQPPGKGLEWIG	EIN	HSGSTNYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	1-1		V_58^{18}	QVQLQQWGAGLLKPSETLSLTCAVYGGSVS	G	YYWS	WIRQPPGKGLEWIG	YIY	YSGSTNNNPSLKS	RATISVDTSKNQFSLNLSSVTAADTAVYCCAR
	3-1	R1	DP-64	QLQLQESGSGLVKPSQTLSLTCAVSGGSIS	SGGYSWS		WIRQPPGKGLEWIG	YIY	HSGSTYYNPSLKS	RVTISVDRSKNQFSLKLSSVTAADTAVYYCAR
	3-1	R0	DP-65	QVQLQESGPGLVKPSQTLSLTCTVSGGSIS	SGGYYWS		WIRQHPGKGLEWIG	YIY	YSGSTYYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	3-1	R1	V_71-2^4/DP-66	QVQLQESGPGLVKPSETLSLTCTVSGGSVS	SGSYYWS		WIRQPPGKGLEWIG	YIY	YSGSTNYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	2-1	R8	DP-67	QVQLQESGPGLVKPSETLSLTCAVSGYSIS	SG	YYWG	WIRQPPGKGLEWIG	SIY	HSGSTYYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	2-1		$V_H^{SP19}/VH-JA^{20}/V_H4.22^{17}$	QVQLQESGPGLVKPSETLSLTCTVSGYSIS	SG	YYWG	WIRQPPGKGLEWIG	SIY	HSGSTYYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	2-1	R13	$V_{12G-1}{}^{18}/1.9II^3/V_H4.13^{17}$/DP-68	QVQLQESGPGLVKPSDTLSLTCAVSGYSIS	SS	NWWG	WIRQPPGKGLEWIG	YIY	YSGSTYYNPSLKS	RVTMSVDTSKNQFSLKLSSVTAVDTAVYYCAR
	2-1		hv4005[21]	QVQLQESGPGLVKPSQTLSLTCAVSGYSIS	SS	NWWG	WIRQPPGKGLEWIG	YIY	YSGSIYYNPSLKS	RVTMSVDTSKNQFSLKLSSVTAVDTAVYYCAR
	2-1	H	V_11^{17}	QVQLQESGPGLVKPSETLSLTCVVSGGSIS	SS	NWWS	WVRQPPGKGLEWIG	EIY	HSGNPNYNPSLKS	RVTISIDKSKNQFSLKLSSVTAADTAVYYCAR
	2-1		$V_H4.17^{17}/V_H4.23^{17}$/DP-69	QVQLQESGPGLVKPSETLSLTCVVSGGSIS	SS	NWWS	WVRQPPGKGLEWIG	EIY	HSGSPNYNPSLKS	RVTISVDKSKNQFSLKLSSVTAADTAVYYCAR
	2-1		$V_79^{18}/V_H4.19^{17}/V_{IV-4b}{}^2$	QVQLQESGPGLVKPPGTLSLTCAVSGGSIS	SS	NWWS	WVRQPPGKGLEWIG	EIY	HSGSTNYNPSLKS	RVTISVDKSKNQFSLKLSSVTAADTAVYCCAR
	2-1	R0	DP-70	QVQLQESGPGLVKPSGTLSLTCAVSGGSIS	SS	NWWS	WVRQPPGKGLEWIG	EIY	HSGSTNYNPSLKS	RVTISVDKSKNQFSLKLSSVTAADTAVYYCAR
	3-1	R0	$V_H4.18^{17}$	QLQLQESGPGLVKPSETLSLTCTVSGGSIS	SSSYYWG		WIRQPPGKGLEWIG	SIY	YSGSTYYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	3-1		V_2-1^{18}	QLQLQESGPGLVKPSETLSLTCTVSGGSIS	SSSYYWG		WIRQPPGKGLEWIG	SIY	YSGSTYYNPSLKS	RVTISVDTSKNHFSLKLSSVTAADTAVYYCAR
	1-1	R1	$V_{IV-4}{}^2$	QVQLQESGPGLVKPSETLSLTCTVSGGSIS	S	YYWS	WIRQPAGKGLEWIG	RIY	TSGSTNYNPSLKS	RVTMSVDTSKNQFSLKLSSVTAADTAVYYCAR
	1-1	R0	$V_H4.11^{17}$/DP-71	QVQLQESGPGLVKPSETLSLTCTVSGGSIS	S	YYWS	WIRQPPGKGLEWIG	YIY	YSGSTNYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	1-1	R5	V_71-4^4	QVQLQESGPGLVKPSETLSLTCTVSGGSVS	S	YYWS	WIRQPPGKGLEWIG	YIY	YSGSTNYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
	1-1		$V_H4.16^{17}$	QVQLQESGPGLVKPSDTLSLTCTVSGGSIS	S	YYWS	WIRQPPGKGLEWIG	YIY	YSGSTNYNPSLKS	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR
V_H5	1-2	R0	VH251[22]/DP-73	EVQLVQSGAEVKKPGESLKISCKGSGYSFT	S	YWIG	WVRQMPGKGLEWMG	IIYP	GDSDTRYSPSFQG	QVTISADKSISTAYLQWSSLKASDTAMYYCAR
	1-2		V_HVJB^{17}	EVQLVQSGAEVKKPGESLKISCKGSGYSFT	S	YWIG	WVRQMPGKGLEWMG	IIYP	GDSDTRYSPSFQG	QVTISADKPISTAYLQWSSLKASDTAMYYCAR
	1-2	R6	V_HVCW^{17}	EVQLVQSGAEVKKPGESLKISCKGSGYSFT	S	YWIG	WVRQMPGKGLEWMG	IIYP	GDSDTRYSPSFQG	QVTISADKSISTAYLQWSSLKASDTAMYYCAR
	1-2	H,N	$1-v^3$	EVQLLQSAAEVKRPGESLRISCKTSGYSFT	S	YWIH	WVRQMPGKELEWMG	SIYP	GNSDTRYSPSFQG	HVTISADSSSSTAYLQWSSLKASDAAMYYCVR
	1-2		VH32[23]	EVQLVQSGAEVKKPGESLRISCKGSGYSFT	S	YWIS	WVRQMPGKGLEWMG	RIDP	SDSYTNYSPSFQG	HVTISADKSISTAYLQWSSLKASDTAMYYCAR
	1-2	R0	V_HVRG^{17}/V_HVMW^{17}	EVQLVQSGAEVKKPGESLRISCKGSGYSFT	S	YWIS	WVRQMPGKGLEWMG	RIDP	SDSYTNYSPSFQG	HVTISADKSISTAYLQWSSLKASDTAMYYCAR
V_H6	3-5	R0	$V_H-VI^{24}/6-1G1^3$/DP-74	QVQLQQSGPGLVKPSQTLSLTCAISGDSVS	SNSAAWN		WIRQSPSRGLEWLG	RTYYR	SKWYNDYAVSVKS	RITINPDTSKNQFSLQLNSVTPEDTAVYYCAR

Fig. 2(a) *continued*

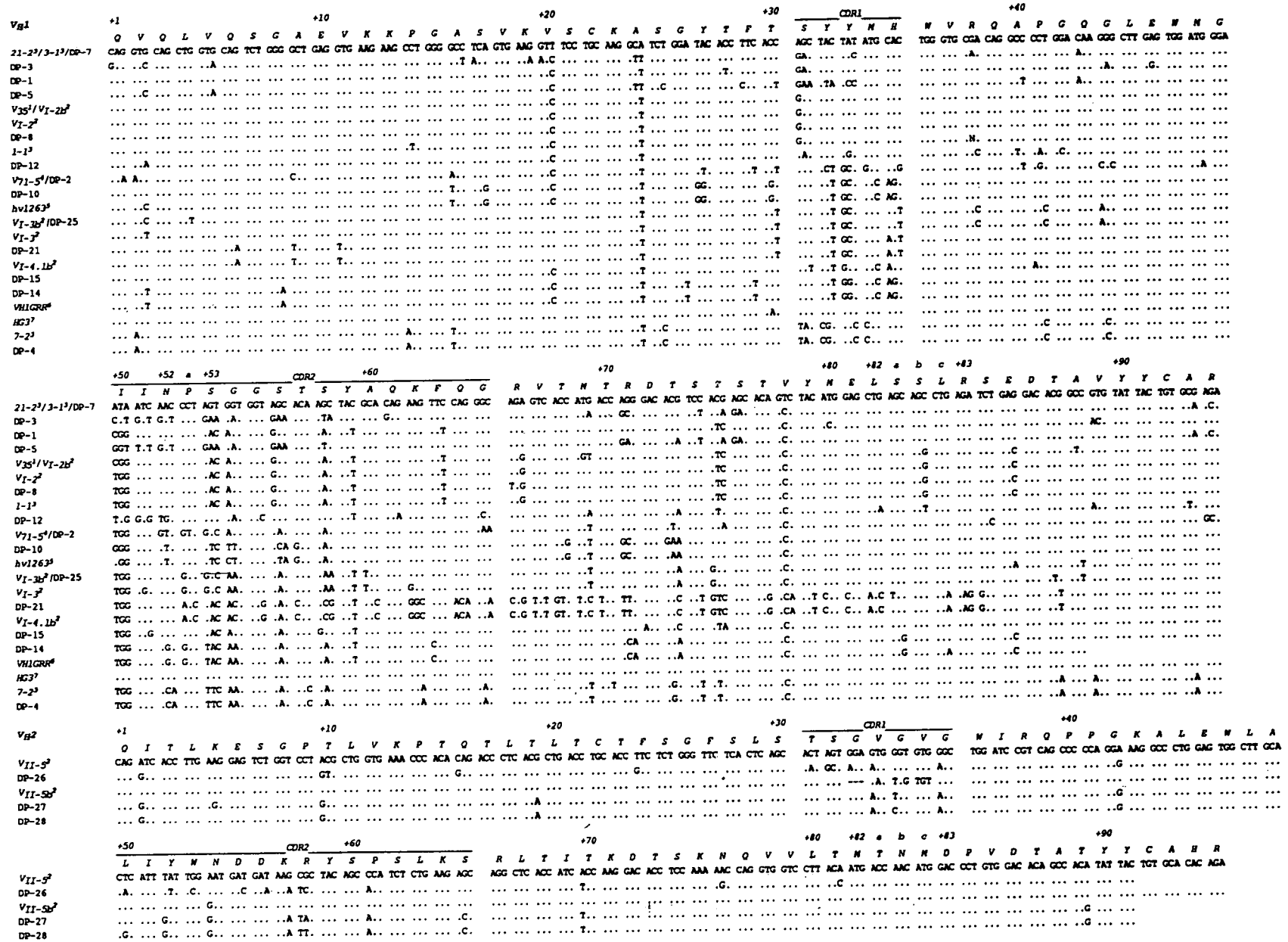


Fig. 2(b)

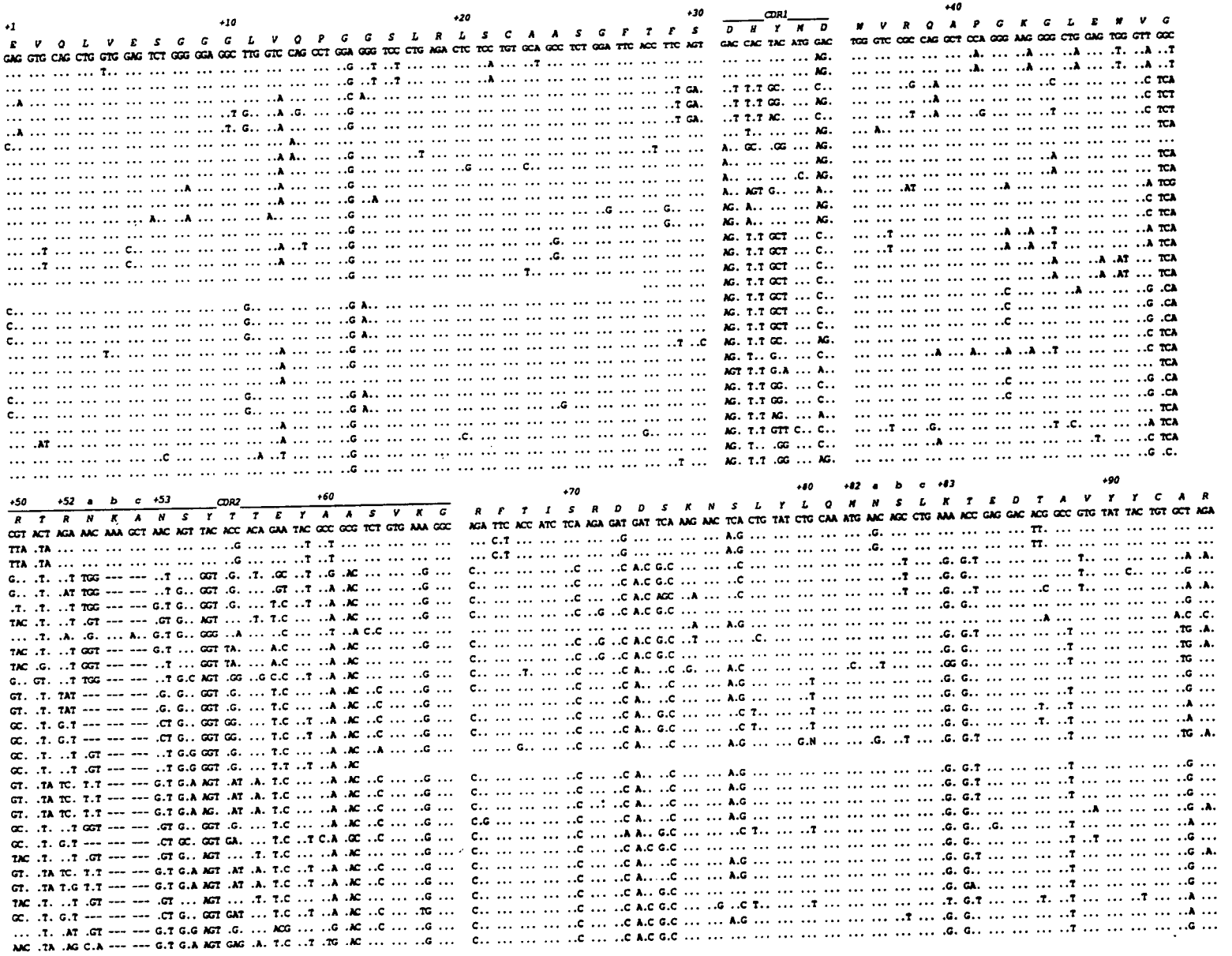


Fig. 2(b) *continued*

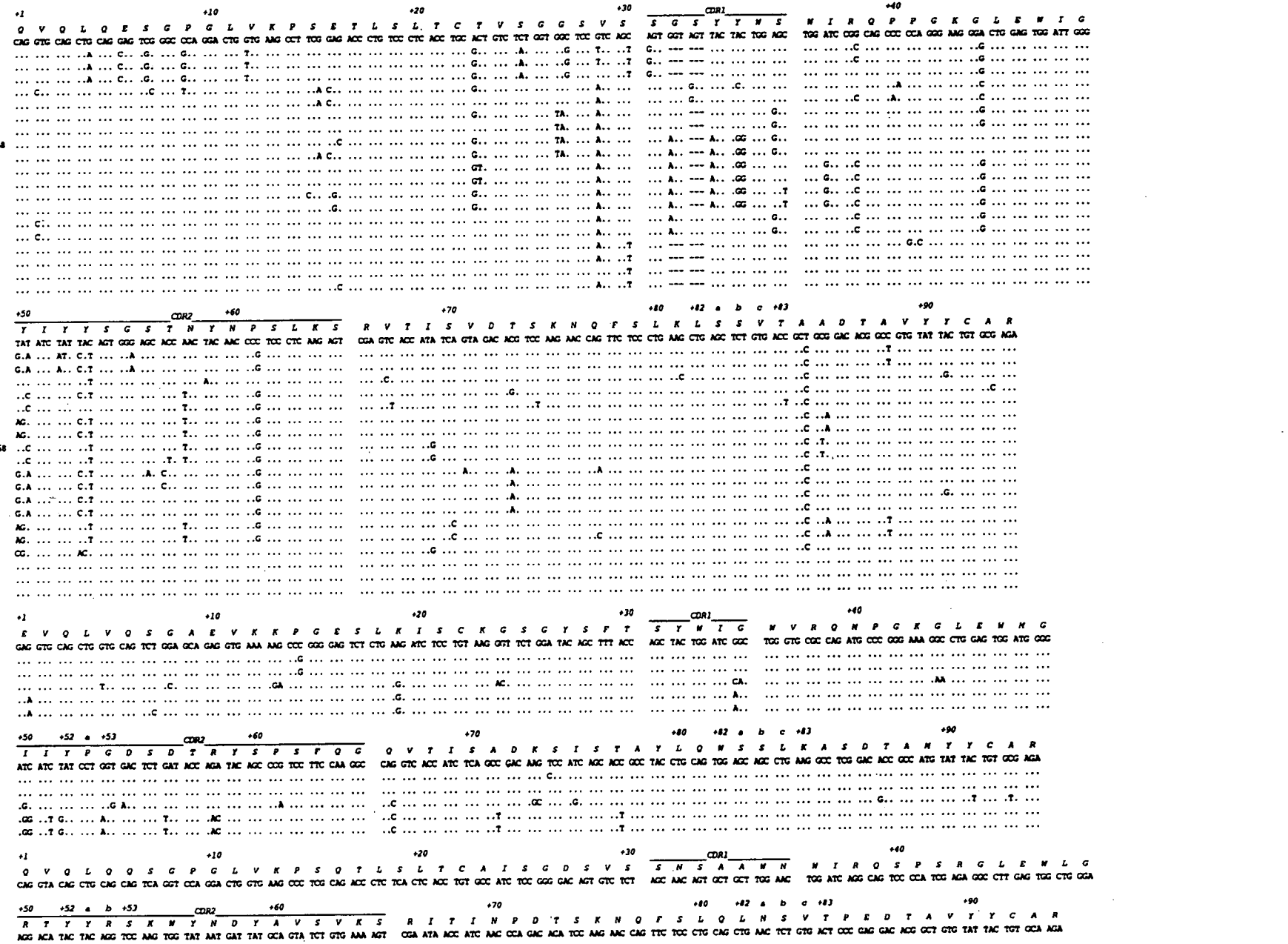


Fig. 2(b) *continued*

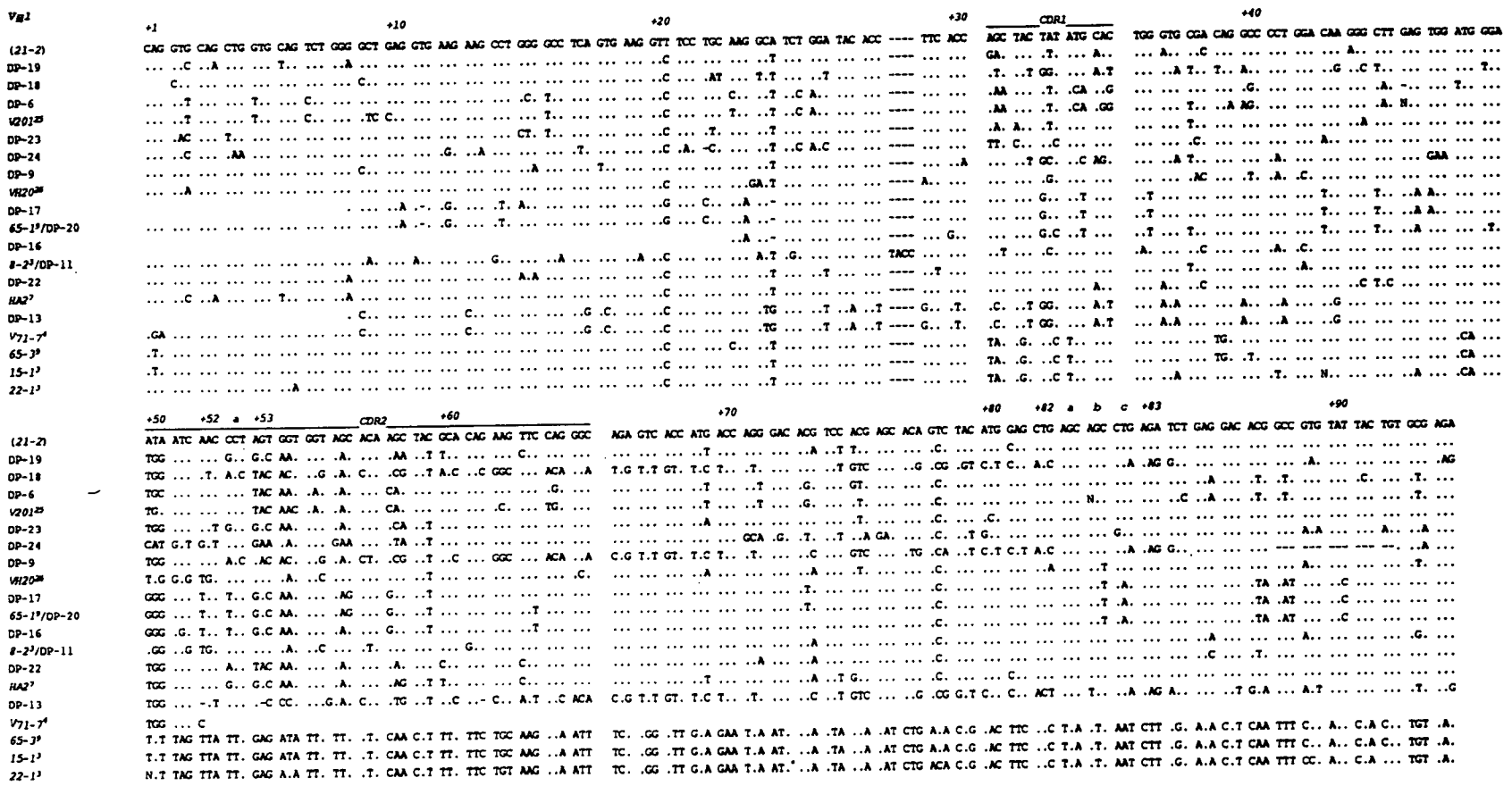

VH1
+1
+10
+20
+30
CDR1
+40
(21-2)
DP-19
DP-18
DP-6
VI201[25]
DP-23
DP-24
DP-9
VH20[26]
DP-17
65-1[3]/DP-20
DP-16
8-2[3]/DP-11
DP-22
HA2[7]
DP-13
V71-7[4]
65-3[3]
15-1[3]
22-1[3]
+50
+52
a
+53
CDR2
+60
+70
+80
+82
a
b
c
+83
+90

Fig. 2(c)

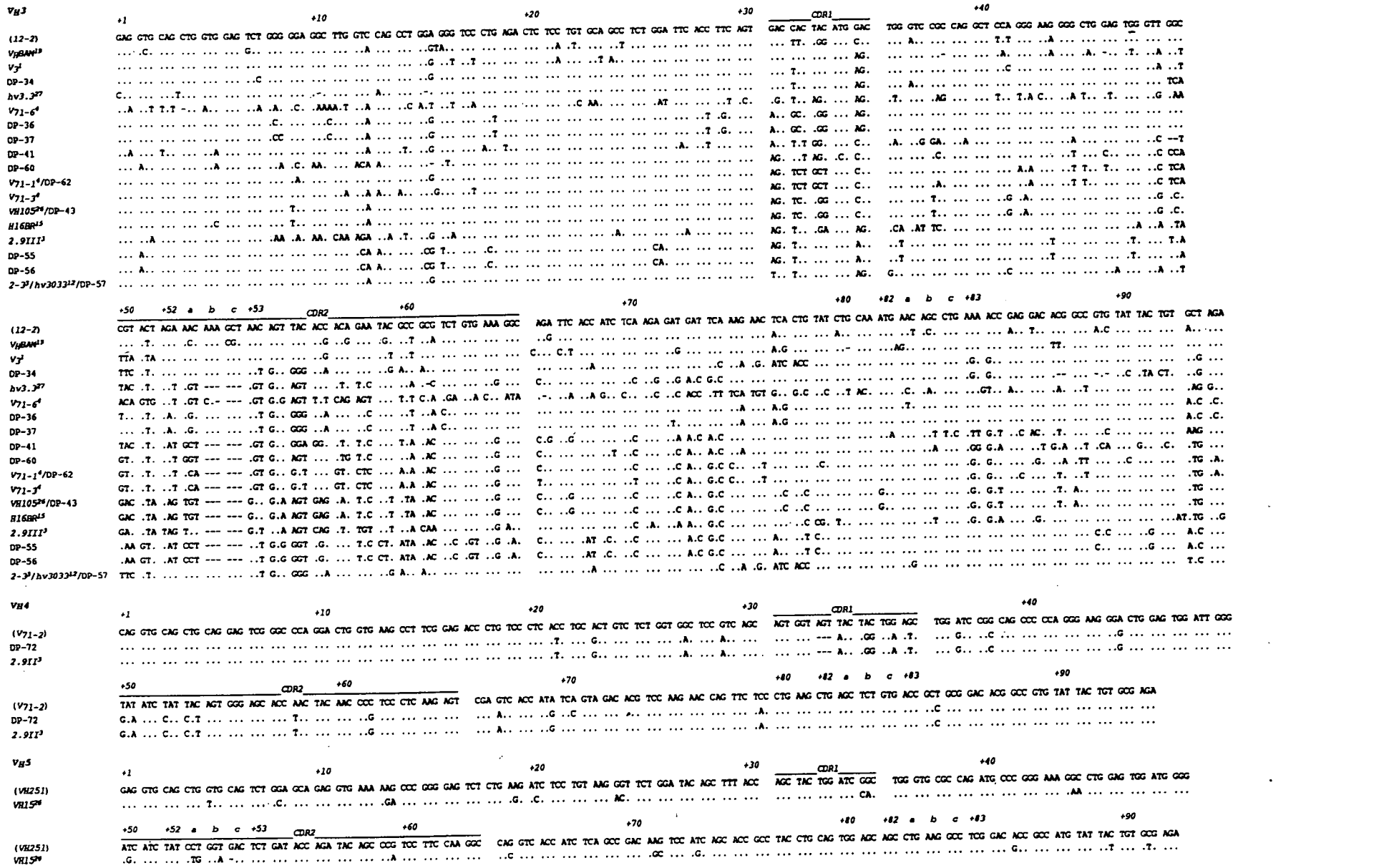
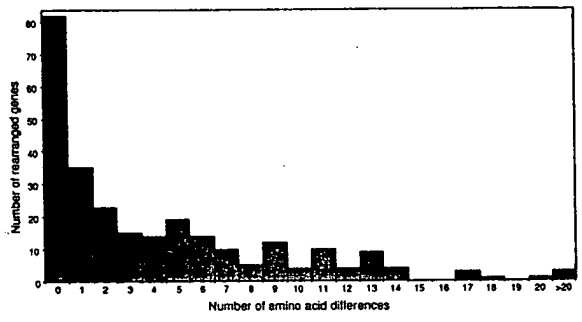


Fig. 2(c) *continued*

Figure 2. Directory of germline V_H segments. Genes are divided into their respective families and the framework (FR) and complementarity-determining regions (CDR) are as defined by Kabat *et al.* (1991). Where 2 genes have identical nucleotide sequences, both are shown separated by a slash. Genes prefixed DP are from this study. Previously published genes are shown in italics and suffixed according to source: [1]Matsuda *et al.* (1988); [2]Shin *et al.* (1991); [3]Berman *et al.* (1988); [4]Kodaira *et al.* (1986); [5]Chen *et al.* (1989); [6]Friedman *et al.* (1991); [7]Rechavi *et al.* (1983); [8]Buluwella *et al.* (1988); [9]Matsuda *et al.* (1990); [10]Baer *et al.* (1988); [11]Olee *et al.* (1991); [12]Chen (1990); [13]Pascual *et al.* (1990); [14]Mathyssens & Rabbitts (1980), corrected by Chen *et al.* (1988); [15]Rechavi *et al.* (1982); [16]van Es *et al.* (1991); [17]Sanz *et al.* (1989c); [18]Lee *et al.* (1987); [19]Baer *et al.* (1985); [20]Denny *et al.* (1986); [21]Chen & Yang (1990); [22]Shen *et al.* (1987), corrected by Sanz *et al.* (1989c); [23]Humphries *et al.* (1988), corrected by Sanz *et al.* (1989c); [24]Buluwela & Rabbitts (1988); [25]Takahashi *et al.* (1984); [26]Humphries *et al.* (1988); [27]Turnbull *et al.* (1987). A number of published V_H segments have discrepancies between the sequence described in the original literature and that submitted to GenBank. We have used the sequence from the literature. (a) Protein sequences of V_H segments with open reading frames. Sequences are shown in single-letter amino acid code and have been aligned according to Kabat *et al.* (1991), except in CDR1. where padding is according to the H1 loop structure (Chothia *et al.*, 1992). Sequences are arranged alphabetically by CDR1 and, where these are identical, by CDR2. Sequences that have identical translated CDRs 1 and 2 are grouped with brackets. The canonical structure classes of H1 (CDR1) and H2 (CDR2) are shown, separated by a dash (see Chothia *et al.*, 1992). The sequence marked " . " has insufficient data to be classified. Features: *X*, may be defective on structural grounds (see Chothia *et al.*, 1992); *H*. heptamer differs from the conserved 5′-CACAGTG-3′ motif; *N*, nonamer sequence differs from the family consensus; *R*, seen as rearranged gene (see Fig. 3 and Results), the smallest number of amino acid differences between each rearranged gene and its closest germline counterpart is shown. Thus, V_{I-2} and DP-8 have identical encoded CDRs 1 and 2, have H1 canonical structure class 1, H2 canonical structure class 3 and have both been seen as rearranged genes with no amino acid differences. (b) DNA sequences of V_H segments with open reading frames. Sequences have been aligned to a master sequence (see Results) and nucleotides identical to this sequence are shown as dots and deletions are indicated by a dash. Numbering is according to the corresponding amino acid residue. (c) V_H pseudogenes. Sequences of germline V_H segments with either frame shifts or stop codons, which are therefore considered to be pseudogenes. Because of their different lengths and heterogeneity, they have been aligned to the master sequence (which for each family is the same as in (b), above) by a maximum alignment program (MacVector). Nucleotides that are identical to the master sequence are shown as dots and deletions are indicated by a dash. Insertions have been placed in between adjacent codons.

Germline Segment	Rearranged Gene	Reference	Number of Amino Acid Changes
DP-3	31P19	(1)	5
$V_{H}2$	16-6	(2)	0
DP-8	15-6	(2)	0
DP-10	AND	(3)	0
h×1263	III-2R	(4)	3
$V_{H}3b$	VE3D10	(5)	0
$V_{H}4.1b$	RF-TS3	(6)	1
DP-15	216B	(7)	2
DP-14	12B	(7)	1
VH1GRR	LS7	(8)	4
21-2	MO30	(9)	3
DP-26	26E	(7)	0
DP-27	M60	(10)	3
DP-28	26D	(7)	0
$V_{H}5$	26A	(7)	0
DP-31	C6H	(11)	0
DP-32	6M9	(12)	2
DP-33	6A1	(13)	4
22-2B	H2F	(4)	1
9-1	M26	(10)	0
DP-42	60P2	(14)	1
8-1B	14J	(7)	1
π1-p1	12H	(7)	10
h×3005f3	M72	(10)	0
GL-SI2	A39	(15)	0
VH26	12C	(7)	0
13-2	36-1	(2)	0
DP-38	21SD	(7)	0
1.9III	αBSA3	(16)	0
3019h9	5A10	(17)	0
DP-51	11O9-10B4	(15)	3
H11	21SL	(7)	2
DP-54	X31	(18)	0
Tou-VH4.21	Fom-1	(19)	2
$V_{H}5$	21SH	(7)	0
DP-64	21SO	(7)	1
DP-65	14L	(7)	0
$V_{H}1$-2	216H	(7)	1
DP-67	A45S	(20)	8
$V_{H}2O$-1	Ab26	(21)	13
DP-70	21SC	(7)	0
$V_{H}4.1B$	14D	(7)	0
$V_{H}V$-4	21SA	(7)	1
$V_{H}4.11$	7-2	(22)	0
$V_{H}1$-6	Pag-1	(23)	5
VH2S1	28-3	(2)	0
$V_{H}VCW$	II-1	(4)	6
$V_{H}VRG$	M13	(20)	0
V_{H}-VI	17-2	(2)	0

(a)



(b)

Figure 3. Assignment of rearranged human V_H genes to their closest germline counterparts. (a) Germline V_H segments and the closest rearranged V_H gene, references are (1) Bridges *et al.* (1991)†; (2) Deane & Norton (1990)‡§; (3) Kipps *et al.* (1989)§; (4) Manheimer-Lory *et al.* (1991)†; (5) Noma *et al.* (1984); (6) Pascual *et al.* (1990)†; (7) Marks *et al.* (1991*b*)‡; (8) Silberstein *et al.* (1989)†; (9) Larrick *et al.* (1989*a*); (10) Schroeder & Wang (1990)¶; (11) Ermel *et al.* (1991)†; (12) Karr *et al.* (1991)¶; (13) Brown *et al.* (1991)†; (14) Schroeder *et al.* (1987)¶; (15) Geng *et al.* (1991)†; (16) Marks *et al.* (1991*a*)‡; (17) see Olee *et al.* (1991)¶; (18) Timmers *et al.* (1991); (19) Bye *et al.* (1992)‡; (20) Schutte *et al.* (1991)¶†§; (21) Sanz *et al.* (1989*a*)†; (22) Desai *et al.* (1990)§; (23) Hughes-Jones *et al.* (1990). (b) Distribution of the number of amino acid differences between each rearranged V_H gene (268 examples) and its closest germline counterpart. Data were taken from the above references and Kenten *et al.* (1982); Takahashi *et al.* (1984); Kudo *et al.* (1985); Mensink *et al.* (1986); Dersimonian *et al.* (1987)†; Shen *et al.* (1987)§; Berman *et al.* (1988); Meeker *et al.* (1988)§; Newkirk *et al.* (1988); Cairns *et al.* (1989)†; Carroll *et al.* (1989); Chen *et al.* (1989)§; Dersimonian *et al.* (1989)†; Gillies *et al.* (1989); Kishimoto *et al.* (1989); Larrick *et al.* (1989*b*); Logtenberg *et al.* (1989)¶†; Nakatani *et al.* (1989); Nickerson *et al.* (1989)¶; Sanz *et al.* (1989*b*)†; Yasui *et al.* (1989); Akahori *et al.* (1990); Felgenhauer *et al.* (1990); Friedländer *et al.* (1990); Guillaume *et al.* (1990)¶†; Robbins *et al.* (1990)†; Roudier *et al.* (1990)†§; Siminovitch & Chen (1990)†; Spatz *et al.* (1990)§; van der Heijden *et al.* (1990); White *et al.* (1990); Andris *et al.* (1990); Ezaki *et al.* (1991)†; Friedman *et al.* (1991)†; Kuppers *et al.* (1991)§; Mortari *et al.* (1991); Pascual *et al.* (1991); Rioux *et al.* (1991)†; Silberstein *et al.* (1991); van Es *et al.* (1991)†; Mierau *et al.* (1992) . Some of the references include sequences from several different rearranged V_H genes: all of the sequences (except where the genes could not be assigned, see Results) have been used. For key to annotation of references (†, ‡, § and ¶) see Discussion.

family:

(V_H1, 5′-TCAGAAACC-3′;
V_H2, 5′-ACAAAAACC-3′;
V_H3, 5′-ACACAAACC-3′;
V_H4, 5′-ACAAAAACC-3′ or
5′-ACACAAACC-3′;
V_H5, 5′-TCTAAAACC-3′;
V_H6, 5′-ACACAAACC-3′).

Where the nonamer sequence differs from the family consensus the V_H segment is marked *N*.

We compiled a database of 292 rearranged (but not necessarily functional) V_H genes and assigned 268 of these, from 64 different sources (see legend to Fig. 3), to their closest germline counterparts. In Figure 3(a) we list the V_H segments, each with an example of a rearranged V_H gene with the smallest number of amino acid differences. These data are summarized in Figure 2(a), with sequences marked *R* having rearranged counterparts with the indicated number of amino acid differences. The distribution of the number of amino acid differences across all 268 assigned rearranged genes is shown in Figure 3(b): 215 of the 292 rearranged V_H genes in our database have germline counterparts seen in DP (data not shown).

We were unable to assign 24 rearranged genes from the V_H3 (VDJ191, Mensink *et al.* (1986); X51, X61, X71, Timmers *et al.* (1991); K6H6, K4B8, K5B8, K5G5, K6F5, K5C7, Kon *et al.* (1987); 333, 1H1, 2C12, 2A12, 1B11, 112, 115 and 126, Cleary *et al.* (1986)) and V_H4 (TS2, Shen *et al.* (1987); HIVB, Andris *et al.* (1991); C6B2, Hoch & Schwaber (1987); 2A4, Davidson *et al.* (1990); 12-3, 30-2, Deane & Norton (1990)) families. Of these, 12-3 (Deane & Norton, 1990) is almost certainly the result of a PCR cross-over and the others appear to be derived from a possible four to six unknown germline V_H segments.

(e) *Germline sequence variability*

Based on data from Figure 2(a), we have constructed variability plots, shown in Figure 4, for germline V_H segments with open reading frames from all six families, as well as separate plots for the V_H1 and V_H3 families. We only excluded those sequences marked *X* which may be defective on structural grounds (see above). At each position, a variability score was calculated as the number of different amino acids at that position, divided by the percentage frequency of occurrence of the most common amino acid (see Kabat *et al.*, 1991).

4. Discussion

(a) *Cloning and sequencing strategy*

Our strategy for sequencing V_H segments by PCR amplification of genomic DNA is based on the use of

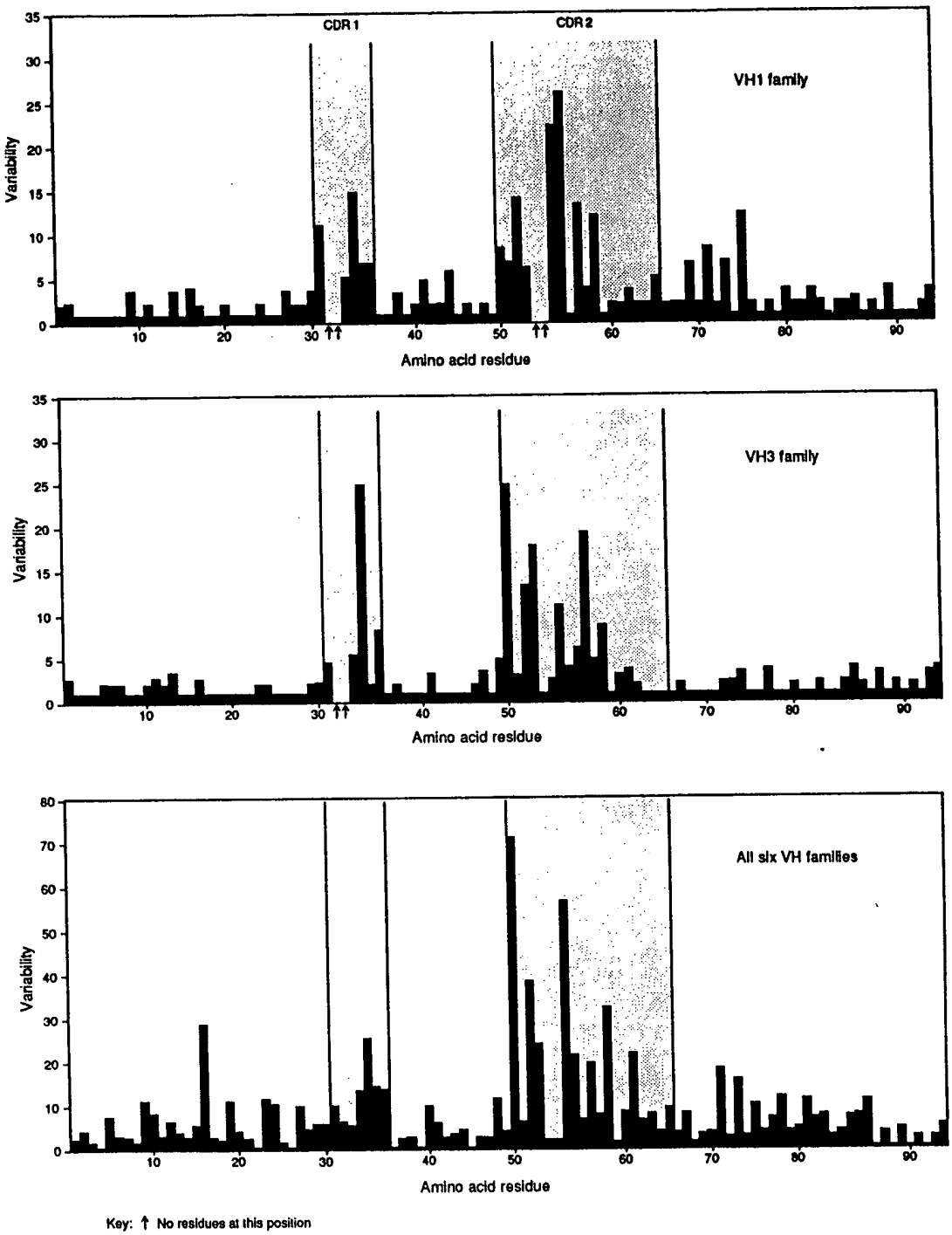


Figure 4. Variability plot for germline V_H segments. Variability was calculated (see Results) across protein sequences shown in Fig. 2(a), but excluding those that are likely to be defective on structural grounds (marked *X*). Plots were produced for the V_H1 family, V_H3 family and across all 6 families.

family-specific primers designed from the sequences of the six *known* V_H families. We were able to assign most of the rearranged V_H genes to germline V_H segments in Figure 2 with few differences in amino acid sequences (Fig. 3(b)), but may have missed V_H segments that are significantly different in the primer regions; for example, we did not find the germline counterparts of the rearranged genes determined by Cleary *et al.* (1986). Indeed, they have been classified as belonging to a new family (V_H7) by some authors (Schroeder *et al.*, 1990), but they might also be highly mutated genes derived from a known germline V_H segment (especially as they were derived from B-cell lymphomas).

Since our aim was to determine the structural repertoire of human V_H segments, the majority of primers were designed to amplify genes with "functional" heptamer recombination sequences

(5′-CACAGTG-3′). We have therefore missed some genes with different heptamers, which presumably includes some pseudogenes. For example, three sequences which were amplified with internal V_H3 primers and have unusual heptamer sequences, DP-59/V_H19, DP-60 and DP-62/$V_{71\text{-}1}$, were not amplified using the heptamer primers. It is, however, unclear what constitutes a functional heptamer; indeed, in a recent study, Shin *et al.* (1991) discovered two V_H2 segments with an unusual heptamer sequence (5′-CACAAAG-3′). One of these segments has been seen as a rearranged gene (see Fig. 3(a)). This suggests that the 5′-CACAGTG-3′ heptamer sequence is not the only one used for recombination and, consequently, that the D segment heptamer may also be degenerate. This, and the fact that these V_H2 segments ($V_{11\text{-}5}$) have an additional amino acid residue in framework 3, may explain the poor performance of our V_H2 primers and the relatively low number of V_H2 segments isolated here.

In addition, those genes with open reading frames (Fig. 2(a) and (b)) may be non-functional for other reasons. For example, the V_H1 segments 1-1 (Berman *et al.*, 1988) and $V_{71\text{-}5}$ (Kodaira *et al.*, 1986) have single base differences in the recombination nonamer and the leader intron splice site, respectively, and 1-v (Berman *et al.*, 1988) has a frame shift in the leader exon. Certain V_H segments may also be defective on structural grounds (marked *X* in Fig. 2(a), see Chothia *et al.*, 1992).

To avoid polymerase copying errors, we screened more than 2000 clones using motif- or gene-specific oligonucleotide probes to ensure identical nucleotide sequences from *two independent* amplifications. Copying errors fell into two categories: base substitutions and PCR cross-overs. Substitutions might have been reduced by using a polymerase with a 5′ to 3′ proof-reading activity such as Vent (New England Biolabs, Beverly, MA) or Pfu (Stratagene, La Jolla, CA) DNA polymerases. However, under a range of conditions, these polymerases performed poorly (data not shown). PCR cross-overs occurred within the region of greatest homology, and were most easily detected by unexpected combinations of CDR1 and CDR2 due to a cross-over in framework 2. This emphasizes the importance of confirmation from independent amplifications rather than from multiple clones of the same PCR; indeed, germline V_H segments hv3005b54, hv3019b13, hv3019b18 (Olee *et al.*, 1991) and $V_H4.12$, $V_H4.14$, $V_H4.15$ (Sanz *et al.*, 1989*c*) may be the result of PCR artifacts (see above).

(b) *Polymorphism*

In our directory (Fig. 2), which contains data from many individuals, we have a total of 122 V_H segments with different nucleotide sequences (41 V_H1 segments, 5 V_H2 segments, 46 V_H3 segments, 22 V_H4 segments, 7 V_H5 segments and 1 V_H6 segment), including 83 V_H segments with open reading frames and 39 pseudogenes. However, we cannot exclude polymorphism and allelic variation or distinguish between identical V_H genes at different loci (possibly the result of a recent duplication).

Southern blot analyses of restriction digests of DNA using cDNA probes (van Dijk *et al.*, 1991), germline coding and flanking region probes (Souroujon *et al.*, 1989) or short sequence-specific probes (Sanz *et al.*, 1989*c*; Sasso *et al.*, 1990; van Dijk *et al.*, 1991) have demonstrated restriction fragment length polymorphisms (RFLPs) in the V_H3, V_H4 and V_H5 families. Some insertion/deletion polymorphisms have also been characterized and shown to involve, for example, at least one V_H2, one V_H3 and one V_H5 gene (Chen & Yang, 1990; Walter *et al.*, 1990), and one V_H1 gene (Shin *et al.*, 1991). Indeed, we failed to clone from DP several V_H segments reported in the literature, despite using suitable PCR primers and probes. Some of the V_H segments not amplified from DP are also missing in other individuals. For example, of the V_H4 segments not amplified from our donor, one (V_{58}) seen in a Japanese study (Lee *et al.*, 1987) was not found in an American study (Sanz *et al.*, 1989*c*) and the absence of a second V_H5 segment, VH32 (see Sanz *et al.*, 1989*c*), from our donor may be due to a deletion polymorphism affecting V_H5 genes in 50% of individuals (Sam *et al.*, 1988).

In our directory, we found that the nucleotide sequences of 23 V_H segments from DP with open reading frames were identical to those from unrelated individuals. We found other V_H segments with a few nucleotide differences but with identical translated CDRs 1 and 2 (bracketed in Fig. 2(a)) and these may correspond to different alleles. Thus, the following V_H1 segments differ by one nucleotide: $V_{I\text{-}2}$, DP-8 and 1-1; DP-21 and $V_{I\text{-}4\text{-}1b}$; DP-14 and VH1GRR; 21-2/3-1/DP-7 and HG3; 7-2 and DP-4. The following V_H3 segments differ by one to six nucleotides: VHD26 and DP-30; DP-42 and 8-1B; 65-2/DP-44 and DP-45; f1-p1 and DP-61; hv3005, hv3005f3 and GL-SJ2/DP-46. The following V_H4 segments differ by one or two nucleotides: DP-67 and V_H^{SP}/VH-JA/$V_H4.22$; V_{79}/$V_H4.19$/$V_{IV\text{-}4b}$ and DP-70; $V_H4.18$ and $V_{2\text{-}1}$; $V_H4.11$/DP-71, $V_{71\text{-}4}$ and $V_H4.16$. The following V_H5 segments differ by one or two nucleotides: VH251/DP-73, V_HVJB and V_HVCW; VH32 and V_HVRG/V_HVMW. Of course, other V_H segments, for example, DP-10 and hv1263, and $V_{I\text{-}3b}$/DP-25 and $V_{I\text{-}3}$ may also be alleles, but they encode differences in the CDRs and have therefore been grouped separately. This is consistent with the suggestion that even diverse V_H segments ($V_{II\text{-}5}$ and $V_{II\text{-}5b}$; $V_{IV\text{-}4}$ and $V_{IV\text{-}4b}$) could be alleles (Shin *et al.*, 1991).

Hence, we find a "core" of V_H segments with open reading frames that are highly conserved in the antigen binding regions and differ by only a small number of nucleotides in the framework regions. This limited sequence polymorphism between unrelated individuals together with the insertion/deletion polymorphism agrees with the suggestion that the germline V_H repertoire is derived from a population of diverse haplotypes with a small

number of alleles at each locus (Sasso *et al.*, 1990; van Dijk *et al.*, 1991).

In contrast to the limited sequence polymorphism in V_H segments with open reading frames, only five pseudogenes amplified from DP are identical to V_H segments seen in unrelated individuals and a further five pairs differ by one or two nucleotides. The finding that certain pseudogenes are identical, or are very similar, in unrelated individuals (see Fig. 2(c)) has been previously noted (Kodaira *et al.*, 1986) and may indicate a physiological role for them, possibly as donors for gene conversion, as in the chicken (Reynaud *et al.*, 1989).

(c) *Assignment of rearranged genes*

As shown in Figure 3(b), the majority of rearranged genes, usually derived from mRNA, are very closely related to their germline counterparts. This confirms that these germline genes can be rearranged and transcribed and are probably translated into protein. Some of the differences between the rearranged and germline genes could be due to germline polymorphism, but as this is limited (see above), the majority are probably caused by somatic mutation. In a few examples, the sequences of the rearranged V_H genes appear to be a composite of two V_H segments (215B and 216G; Marks *et al.*, 1991*b*), which presumably arose by PCR cross-over.

The assignment of rearranged human V_H genes to their germline counterparts may help in dissecting mechanisms of the human immune system. It enables us to determine the relative usage of particular V_H segments (the possible underexpression of V_H1 segments and overexpression of V_H4 segments) and the number and location of somatic mutations by which a particular antibody has been shaped. It also allows us to differentiate between immune responses that utilize V_H segments with different levels of somatic mutation. For example, it has been repeatedly suggested that foetal antibodies and autoantibodies are dominated by rarely mutated or unmutated germline V_H genes and that these antibodies are often polyreactive (see Chen *et al.*, 1990; Hillson & Perlmutter, 1990; Siminovitch & Chen, 1990; Pascual & Capra, 1991).

Using our database of human rearranged V_H genes we find that about three-quarters of the genes of foetal origin are germline at the level of amino acid sequence and the rest have no more than five amino acid changes (see references marked ¶ in Fig. 3 legend). However, in the case of autoantibodies (autoimmunity related V_H genes, see references marked † in Fig. 3 legend) there is no clear difference in the overall number of amino acid changes compared to rearranged V_H genes found in normal peripheral blood lymphycytes (see references marked ‡ in Fig. 3 legend). This does not support the concept that autoantibodies are mainly encoded by rarely mutated or unmutated V_H genes and reflects the current uncertainty about the origin of autoantibodies and the role of antigen stimulation (Dersimonian *et al.*, 1990). Other interesting features emerge for different B cell malignancies. Whereas most of the V_H genes isolated from acute lymphoblastic leukaemia (ALL) patients are rarely mutated or unmutated (Berman *et al.*, 1988; Carroll *et al.*, 1989; Deane & Norton, 1990), about half the V_H genes isolated from patients with chronic lymphocytic leukaemia (CLL) contain more than six amino acid changes (see references marked § in Fig. 3 legend). Very highly mutated V_H genes (17, 20, 43 amino acid changes) have been detected in other B cell tumours, such as myelomas (White *et al.*, 1990; Kenten *et al.*, 1982; Yasui *et al.*, 1989).

(d) *Number of human V_H segments*

Estimates of the number of human V_H segments per individual have been based on restriction digests of genomic DNA probed for each V_H family, but are likely to be underestimates (due to bands co-migrating on the gel). For example, Southern blot analyses of digested DNA from HeLa and LA-N-5 cell lines yielded 60 to 80 hybridizing fragments (Berman *et al.*, 1988) but the authors estimated the total number of V_H segments to be between 100 and 200. More recently, two-dimensional pulse field gel electrophoresis of digested homozygous DNA (Walter *et al.*, 1990) suggested a total of 76 V_H segments (25 V_H1 segments, 5 V_H2 segments, 28 V_H3 segments, 14 V_H4 segments, 3 V_H5 segments and 1 V_H6 segment).

We have cloned and sequenced 74 human V_H segments (25 V_H1 segments, 3 V_H2 segments, 34 V_H3 segments, 10 V_H4 segments, 1 V_H5 segment and 1 V_H6 segment). Fifty-one of these have open reading frames, and 23 contain either frame shifts or stop codons and are therefore considered to be pseudogenes. While the number of pseudogenes amplified from DP is likely to be an underestimate due to primer bias, the number of V_H segments with open reading frames (51) seems to correspond to the coding repertoire of an individual. Indeed, 215 of 292 rearranged V_H genes from different (non-DP) individuals have germline counterparts seen in DP. The extent to which our individual is representative of the human population as a whole depends on the *exact* nature of polymorphism within the V_H locus. To determine this, we need a physical map of the V_H segments from individuals with different genetic backgrounds, in which individual V_H loci have been sequenced. This would tell us the number of different sequences in the human V_H segment pool, the total number of loci and the number of alleles at each locus.

(e) *Structural diversity of human germline V_H segments*

In order to focus on the structural diversity of antigen binding sites implicit in the germline V_H repertoire of the human population, we grouped together (bracketed in Fig. 2(a)) those V_H segments that encode identical CDRs 1 and 2. We have selected those V_H segments with rearranged counter-

parts (marked *R* in Fig. 2(a)) and excluded a few V_H segments (marked *X* in Fig. 2(a)), which appear to be defective on structural grounds (Chothia *et al.*, 1992) and therefore are unlikely to contribute to the functional V_H repertoire.

This suggests that the structural diversity encoded by human germline V_H segments is determined by a minimum of 43 groups of rearranged V_H segments, each encoding identical CDR loops. This figure is likely to increase as rearranged counterparts of other germline V_H segments in Figure 2(a) are discovered and as a few additional germline segments are determined from different individuals. However, those V_H segments with heptamers other than the 5′-CACAGTG-3′ motif (marked *H* in Fig. 2(a)) and those with nonamers that differ from the family consensus (marked *N* in Fig. 2(a)) may be unable to recombine and hence not be expressed.

In order to determine the possible extent of sequence diversity, our variability plots (Fig. 4) are calculated using sequence data from all germline V_H segments of the 43 structural groups and those germline sequences for which no rearranged counterparts have yet been discovered. The use of germline V_H segments eliminates the effects of somatic mutation and sampling bias present in variability plots of rearranged V_H genes (Kabat *et al.*, 1991).

The plots are consistent with the classification of framework (FR) and complementarity-determining regions (CDR) defined by Kabat *et al.* (1991), but new features do emerge. Firstly, variability is higher in CDR2 than in CDR1. Secondly, the hypervariable region of CDR2 only comprises residues 50 to 58, rather than 50 to 65, with the last seven residues of CDR2 (59 to 65) being highly conserved within each of the six families. Thirdly, in addition to CDR1 and CDR2, we find two regions of unusually high variability across all six families. One of them is residue 16 and the other is centred around residue 73 and corresponds to a loop adjacent to CDR2. The region in framework 3 is particularly variable in the V_H1 family and may function by altering the conformation of CDR2 for antigen binding, or make additional contacts directly with the antigen (like in the case of the light chain FR3 in the D1.3/E255 complex: Bentley *et al.*, 1990). Alternatively, it may interact with an unidentified ligand involved in the biology of the B cell response, for example, a superantigen (Schroeder *et al.*, 1990; Sasso *et al.*, 1991).

(f) *Conclusion*

Our strategy has enabled us to determine the human germline V_H segments with open reading frames from a single individual (DP). The comparison with germline V_H segments from other individuals and with 292 rearranged V_H genes suggests that sequence polymorphism is limited, and that the directory could be used to map the V_H locus in different individuals, to determine the usage of specific V_H segments in immune responses and to detect somatic mutation or gene conversion events *in vivo*.

The directory indicates that the structural diversity of the germline repertoire for antigen binding is fixed by about 50 groups of V_H segments. Each group encodes identical hypervariable loops and has been seen as a rearranged gene. The limited diversity encoded by germline V_H segments emphasizes the importance of the additional diversity provided by the D and J_H segments and by somatic mutation. It suggests that our repertoire of V_H segments from DP should be sufficient for building libraries of human antibodies with known components (Winter & Milstein, 1991; Marks *et al.*, 1991*a*).

I.M.T. and G. Walter are supported by the Medical Research Council Human Genome Mapping Project, J.D.M. by the Medical Research Council Aids Directed Programme and M.B.L. by a Medical Research Council/Celltech Ltd. training fellowship. We thank D. J. Perry for his many blood donations and his help with genomic DNA preparation.

References

Akahori, Y., Kurosawa, Y., Kamachi, Y., Torii, S. & Matsuoka, H. (1990). Presence of immunoglobulin (Ig) M and IgG double isotype-bearing cells and defect of switch recombination in hyper IgM immunodeficiency. *J. Clin. Invest.* **85**, 1722–1727.

Alzari, P. M., Spinelli, S., Mariuzza, R. A., Boulot, G., Poljak, R. J., Jarvis, J. M. & Milstein, C. (1990). Three-dimensional structure determination of an anti-2-phenyloxazolone antibody: the role of somatic mutation and heavy/light chain pairing in the maturation of an immune response. *EMBO J.* **9**, 3807–3814.

Amit, A. G., Mariuzza, R. A., Phillips, S. E. & Poljak, R. J. (1986). Three-dimensional structure of an antigen–antibody complex at 2·8 Å resolution. *Science*, **233**, 747–753.

Andris, J. S., Johnson, S., Zolla, P. S. & Capra, J. D. (1991). Molecular characterization of five human anti-human immunodeficiency virus type 1 antibody heavy chains reveals extensive somatic mutation typical of an antigen-driven immune response. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 7783–7787.

Baer, R., Chen, K. C., Smith, S. D. & Rabbitts, T. H. (1985). Fusion of an immunoglobulin variable gene and a T cell receptor constant gene in the chromosome 14 inversion associated with T cell tumors. *Cell*, **43**, 705–713.

Baer, R., Forster, A., Lavenir, I. & Rabbitts, T. H. (1988). Immunoglobulin VH genes are transcribed by T cells in association with a new 5′ exon. *J. Exp. Med.* **167**, 2011–2016.

Bentley, G. A., Boulot, G., Riottot, M. M. & Poljak, R. J. (1990). Three-dimensional structure of an idiotope-anti-idiotope complex. *Nature (London)*, **348**, 254–257.

Berek, C. & Milstein, C. (1988). The dynamic nature of the antibody repertoire. *Immunol. Rev.* **105**, 5–26.

Berman, J. E., Mellis, S. J., Pollock, R., Smith, C. L., Suh, H., Heinke, B., Kowal, C., Surti, U., Cantor, C. R. & Alt, F. W. (1988). Content and organization of the human Ig VH locus: definition of three new

VH families and linkage to the Ig CH locus. *EMBO J.* **7**, 727–738.

Borghesi-Nicoletti, C. & Schulze, D. H. (1991). Polymerase chain reaction of genes flanked by short noncontiguous sequence motifs. *Anal. Biochem.* **192**, 449–452.

Bridges, S. L., Lee, S. L., Koopman, W. J. & Schroeder, H. W. (1991). V_H1 gene segments expressed in rheumatoid synovium. *American College of Rheumatology 55th Annual Scientific Meeting*, 17–21 November, Boston.

Brown, C. M. S., Fitzgerald, K. F., Williams, D. G. & Maini, R. N. (1991). Immunoglobulin heavy chain variable region genes expressed by B cells in rheumatoid synovium. *American College of Rheumatology 55th Annual Scientific Meeting*, 17–21 November, Boston.

Buluwela, L. & Rabbitts, T. H. (1988). A VH gene is located within 95 kb of the human immunoglobulin heavy chain constant region genes. *Eur. J. Immunol.* **18**, 1843–1845.

Buluwela, L., Albertson, D. G., Sherrington, P., Rabbits, P. H., Spurr, N. & Rabbitts, T. H. (1988). The use of chromosomal translocations to study human immunoglobulin gene organization: mapping D_H segments within 35 kb of the C_μ gene and identification of a new D_H locus. *EMBO J.* **7**, 2003–2010.

Buluwela, L., Forster, A., Boehm, T. & Rabbitts, T. H. (1989). A rapid procedure for colony screening using nylon filters. *Nucl. Acids Res.* **17**, 452.

Bye, J. M., Carter, C., Cui, Y., Gorick, B. D., Songsivilai, S., Winter, G., Hughes-Jones, N. C. & Marks, J. D. (1992). *J. Clin. Invest.* in the press.

Cairns, E., Kwong, P. C., Misener, V., Ip, P., Bell, D. A. & Siminovitch, K. A. (1989). Analysis of variable region genes encoding a human anti-DNA antibody of normal origin. Implications for the molecular basis of human autoimmune responses. *J. Immunol.* **143**, 685–691.

Carroll, W. L., Yu, M., Lin, M. P. & Korsmeyer, S. J. (1989). Absence of Ig V region gene somatic hypermutation in advanced Burkitt's lymphoma. *J. Immunol.* **143**, 692–698.

Carter, P., Bedouelle, H. & Winter, G. (1985). Improved oligonucleotide site-directed mutagenesis using M13 vectors. *Nucl. Acids Res.* **13**, 4431–4443.

Chen, P. P. (1990). Structural analyses of human developmentally regulated Vh3 genes. *Scand. J. Immunol.* **31**, 257–267.

Chen, P. P. & Yang, P. M. (1990). A segment of human Vh gene locus is duplicated. *Scand. J. Immunol.* **31**, 593–599.

Chen, P. P., Liu, M. F., Sinha, S. & Carson, D. A. (1988). A 16/6 idiotype-positive anti-DNA antibody is encoded by a conserved VH gene with no somatic mutation. *Arthritis Rheum.* **31**, 1429–1431.

Chen, P. P., Liu, M. F., Glass, C. A., Sinha, S., Kipps, T. J. & Carson, D. A. (1989). Characterization of two immunoglobulin VH genes that are homologous to human rheumatoid factors. *Arthritis Rheum.* **32**, 72–76.

Chen, P. P., Olsen, N. J., Yang, P.-M., Soto-Gil, R. W., Olee, T., Siminovitch, K. A. & Carson, D. A. (1990). From human autoantibodies to the fetal antibody repertoire to B cell malignancy: its a small world after all. *Int. Rev. Immunol.* **5**, 239–251.

Chothia, C. & Lesk, A. M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.* **196**, 901–917.

Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith, G. S., Air, G., Sheriff, S., Padlan. E. A., Davies, D., Tulip, W. R., *et al.* (1989). Conformations of immunoglobulin hypervariable regions. *Nature (London)*, **342**, 877–883.

Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992). The structural repertoire of human V_H segments. *J. Mol. Biol.* **227**, 799–817.

Cleary, M. L., Meeker, T. C., Levy, S., Lee, E., Trela, M., Sklar, J. & Levy, R. (1986). Clustering of extensive somatic mutations in the variable region of an immunoglobulin heavy chain gene from a human B cell lymphoma. *Cell*, **44**, 97–106.

Croce, C. M., Shander, M., Martinis, J., Cicurel, L., D'Ancona, G. G., Dolby, T. W. & Koprowski, H. (1979). Chromosomal location of the genes for human immunoglobulin heavy chains. *Proc. Nat. Acad. Sci., U.S.A.* **76**, 3416–3419.

Davidson, A., Manheimer, L. A., Aranow, C., Peterson, R., Hannigan, N. & Diamond, B. (1990). Molecular characterization of a somatically mutated anti-DNA antibody bearing two systemic lupus erythematosus-related idiotypes. *J. Clin. Invest.* **85**, 1401–1409.

Deane, M. & Norton, J. D. (1990). Immunoglobulin heavy chain variable region family usage is independent of tumor cell phenotype in human B lineage leukemias. *Eur. J. Immunol.* **20**, 2209–2217.

Denny, C. T., Yoshikai, Y., Mak, T. W., Smith, S. D., Hollis, G. F. & Kirsch, I. R. (1986). A chromosome 14 inversion in a T-cell lymphoma is caused by site-specific recombination between immunoglobulin and T-cell receptor loci. *Nature (London)*, **320**, 549–551.

Dersimonian, H., Schwartz, R. S., Barrett, K. J. & Stollar, B. D. (1987). Relationship of human variable region heavy chain germ-line genes to genes encoding anti-DNA autoantibodies. *J. Immunol.* **139**, 2496–2501.

Dersimonian, H., McAdam, K. P., Mackworth, Y. C. & Stollar, B. D. (1989). The recurrent expression of variable region segments in human IgM anti-DNA autoantibodies. *J. Immunol.* **142**, 4027–4033.

Dersimonian, H., Long, A., Rubinstein, D., Stollar, B. D. & Schwartz, R. S. (1990). V_H genes of human autoantibodies. *Int. Rev. Immunol.* **5**, 253–264.

Desai, R., Spatz, L., Matsuda, T., Ilyas, A. A., Berman, J. E., Alt, F. W., Kabat, E. A. & Latov, N. (1990). Molecular cloning of a human immunoglobulin heavy chain variable (VH) region with anti-myelin-associated glycoprotein activity. *J. Neuroimmunol.* **26**, 35–41.

Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988). High efficiency transformation of E. coli by high voltage electroporation. *Nucl. Acids Res.* **16**, 6127–6145.

Ermel, R. W., Kenny, T. P. & Robbins, D. L. (1991). Molecular genetic characterization of human monoclonal IgM rheumatoid factors derived from rheumatoid synovial cells. *American College of Rheumatology 55th Annual Scientific Meeting*, 17–21 November, Boston.

Ezaki, I., Kanda, H., Sakai, K., Fukui, N., Shingu, M., Nobunaga, M. & Watanabe, T. (1991). Restricted diversity of the variable region nucleotide sequences of the heavy and light chains of a human rheumatoid factor. *Arthritis Rheum.* **34**, 343–350.

Felgenhauer, M., Kohl, J. & Ruker, F. (1990). Nucleotide sequences of the cDNAs encoding the V-regions of H- and L-chains of a human monoclonal antibody specific to HIV-1-gp41. *Nucl. Acids Res.* **18**, 4927.

Friedlander, R. M., Nussenzweig, M. C. & Leder, P. (1990). Complete nucleotide sequence of the membrane form of the human IgM heavy chain. *Nucl. Acids Res.* **18**, 4278.

Friedman, D. F., Cho, E. A., Goldman, J., Carmack, C. E., Besa, E. C., Hardy, R. R. & Silberstein, L. E. (1991). The role of clonal selection in the pathogenesis of an autoreactive human B cell lymphoma. *J. Exp. Med.* **174**, 525–537.

Frohman, M. A. & Martin, G. R. (1990). Detection of homologous recombinants. In *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J., eds), pp. 228–236, Academic Press, San Diego.

Geng, L., Silverman, E., Isacovics, B., Zheng, L., Dosch, H.-M. & Siminovitch, K. (1991). Characterization and sequence of human monoclonal anti-Ro/La antibodies. *American College of Rheumatology 55th Annual Scientific Meeting*, 17–21 November, Boston.

Gillies, S. D., Dorai, H., Wesolowski, J., Majeau, G., Young, D., Boyd, J., Gardner, J. & James, K. (1989). Expression of human anti-tetanus toxoid antibody in transfected murine myeloma cells. *Bio/Technology*, **7**, 799–804.

Gronenborn, B. (1976). Overproduction of phage lambda repressor under control of the lac promotor of Escherichia coli. *Mol. Gen. Genet.* **148**, 243–250.

Guillaume, T., Rubinstein, D. B., Young, F. Tucker, L., Logtenberg, T., Schwartz, R. S. & Barrett, K. J. (1990). Individual VH genes detected with oligonucleotide probes from the complementarity-determining regions. *J. Immunol.* **145**, 1934–1945.

Hillson, J. L. & Perlmutter, R. M. (1990). Autoantibodies and the fetal antibody repertoire. *Int. Rev. Immunol.* **5**, 215–29.

Hoch, S. & Schwaber, J. (1987). Identification and sequence of the VH gene elements encoding a human anti-DNA antibody. *J. Immunol.* **139**, 1689–1693.

Hughes-Jones, N. C., Bye, J. M., Beale, D. & Coadwell, J. (1990). Nucleotide sequences and three-dimensional modelling of the VH and VL domains of two human monoclonal antibodies specific for the D antigen of the human Rh-blood-group system. *Biochem. J.* **268**, 135–140.

Humphries, C. G., Shen, A., Kuziel, W. A., Capra, J. D., Blattner, F. R. & Tucker, P. W. (1988). A new human immunoglobulin VH family preferentially rearranged in immature B-cell tumours. *Nature (London)*, **311**, 446–449.

Kabat, E. A. & Wu, T. T. (1971). Attempts to locate complementarity-determining residues in the variable positions of light and heavy chains. *Ann. N.Y. Acad. Sci.* **190**, 382–393.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*. 5th Edit. U.S. Department of Health and Human Services, Public Health Service National Institutes of Health.

Karr, N. S., Sasso, E. H., Hillson, J. L., Oppliger, I. R. & Mannik, M. (1991). Interaction between staphylococcal protein A and human antibodies of the VHIII family. *American College of Rheumatology 55th Annual Scientific Meeting*. 17–21 November, Boston.

Kenten, J. H., Molgaard, H. V., Houghton, M., Derbyshire, R. B., Viney, J., Bell, L. O. & Gould, H. J. (1982). Cloning and sequence determination of the gene for the human immunoglobulin epsilon chain expressed in a myeloma cell line. *Proc. Nat. Acad. Sci., U.S.A.* **79**, 6661–6665.

Kipps, T. J., Tomhave, E., Pratt, L. F., Duffy, S., Chen, P. P. & Carson, D. A. (1989). Developmentally restricted immunoglobulin heavy chain variable region gene expressed at high frequency in chronic lymphocytic leukemia. *Proc. Nat. Acad. Sci., U.S.A.* **86**, 5913–5917.

Kirsch, I. R., Morton, C. C., Nakahara, K. & Leder, P. (1982). Human immunoglobulin heavy chain genes map to a region of translocations in malignant B lymphocytes. *Science*, **216**, 301–303.

Kishimoto, T., Okajima, H., Okumoto, T. & Taniguchi, M. (1989). Nucleotide sequences of the cDNAs encoding the V-regions of H- and L-chains of a human monoclonal antibody with broad reactivity to malignant tumor cells. *Nucl. Acids Res.* **17**, 4385.

Kodaira, M., Kinashi, T., Umemura, I., Matsuda, F., Noma, T., Ono, Y. & Honjo, T. (1986). Organization and evolution of variable region genes of the human immunoglobulin heavy chain. *J. Mol. Biol.* **190**, 529–541.

Kon, S., Levy, S. & Levy, R. (1987). Retention of an idiotypic determinant in a human B-cell lymphoma undergoing immunoglobulin variable-region mutation. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 5053–5057.

Kudo, A., Ishihara, T., Nishimura, Y. & Watanabe, T. (1985). A cloned human immunoglobulin heavy chain gene with a novel direct-repeat sequence in 5′ flanking region. *Gene*, **33**, 181–189.

Kuppers, R., Gause, A. & Rajewsky, K. (1991). B cells of chronic lymphatic leukemia express V genes in unmutated form. *Leuk. Res.* **15**, 487–496.

Larrick, J. W., Danielsson, L., Brenner, C. A., Abrahamson, M., Fry, K. E. & Borrebaeck, C. A. (1989*a*). Rapid cloning of rearranged immunoglobulin genes from human hybridoma cells using mixed primers and the polymerase chain reaction. *Biochem. Biophys. Res. Commun.* **160**, 1250–1256.

Larrick, J. W., Danielsson, L., Brenner, C. A., Wallace, E. F., Abrahamson, M., Fry, K. E. & Borrebaeck, C. A. K. (1989*b*). Polymerase chain reaction using mixed primers: cloning of human monoclonal antibody variable region genes from single hybridoma cells. *Bio/Technology*, **7**, 934–938.

Lee, K. H., Matsuda, F., Kinashi, T., Kodaira, M. & Honjo, T. (1987). A novel family of variable region genes of the human immunoglobulin heavy chain. *J. Mol. Biol.* **195**, 761–768.

Logtenberg, T., Young, F. M., Van, E. J., Gmelig, M. F. & Alt, F. W. (1989). Autoantibodies encoded by the most Jh-proximal human immunoglobulin heavy chain variable region gene. *J. Exp. Med.* **170**, 1347–1355.

Manheimer-Lory, A., Katz, J. B., Pillinger, M., Ghossein, C., Smith, A. & Diamond, B. (1991). Molecular characteristics of antibodies bearing an anti-DNA-associated idiotype. *J. Exp. Med.* **174**, 1639–1652.

Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D. & Winter, G. (1991*a*). By-passing immunization: human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.* **222**, 581–597.

Marks, J. D., Tristrem, M., Karpas, A. & Winter, G. (1991*b*). Oligonucleotide primers for polymerase chain reaction amplification of human immunoglobulin variable genes and design of family-specific oligonucleotide probes. *Eur. J. Immunol.* **21**, 985–991.

Maruyama, I. N. (1990). Estimation of errors in the

polymerase chain reaction. *Techniques-J. Meth. Cell. Mol. Biol.* **2**, 302–303.

Matsuda, F., Lee, K. H., Nakai, S., Sato, T., Kodaira, M., Zong, S. Q., Ohno, H., Fukuhara, S. & Honjo, T. (1988). Dispersed localization of D segments in the human immunoglobulin heavy-chain locus. *EMBO J.* **7**, 1047–1051.

Matsuda, F., Shin, E. K., Hirabayashi, Y., Nagaoka, H., Yoshida, M. C., Zong, S. Q. & Honjo, T. (1990). Organization of variable region segments of the human immunoglobulin heavy chain: duplication of the D_5 cluster within the locus and interchromosomal translocation of variable region segments. *EMBO J.* **9**, 2501–2506.

Matthyssens, G. & Rabbitts, T. H. (1980). Structure and multiplicity of genes for the human immnoglobulin heavy chain variable region. *Proc. Nat. Acad. Sci., U.S.A.* **77**, 6561–6565.

Meeker, T. C., Grimaldi, J. C., O'Rourke, R., Loeb, J., Juliusson, G. & Einhorn, S. (1988). Lack of detectable somatic hypermutation in the V region of the Ig H chain gene of a human chronic B lymphocytic leukemia. *J. Immunol.* **141**, 3994–3998.

Mensink, E. J., Schuurman, R. K., Schot, J. D., Thompson, A. & Alt, F. W. (1986). Immunoglobulin heavy chain gene rearrangements in X-linked agammaglobulinemia. *Eur. J. Immunol.* **16**, 963–967.

Mierau, R., Gause, A., Kuppers, R., Michels, M., Mageed, R. A., Jefferis, R. & Genth, E. (1992). A human monoclonal IgA rheumatoid factor using the VkIV light chain gene. *Rheum. Int.* **12**, 23–31.

Miller, J. H. (1972). *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mills, D. R. & Kramer, F. R. (1979). Structure-independent nucleotide sequence analysis. *Proc. Nat. Acad. Sci., U.S.A.* **76**, 2232–2235.

Mortari, F., Ochs, H. D., Wedgwood, R. J. & Schroeder, H. J. (1991). Immunoglobulin variable heavy chain cDNA sequence from a patient with X-linked agammaglobulinemia. *Nucl. Acids Res.* **19**, 673.

Nakatani, T., Nomura, N., Horigome, K., Ohtsuka, H. & Noguchi, H. (1989). Functional expression of human monoclonal antibody genes directed against pseudomonal exotoxin A in mouse myeloma cells. *Bio/Technology*, **7**, 805–810.

Newkirk, M. M., Gram, H., Heinrich, G. F., Ostberg, L., Capra, J. D. & Wasserman, R. L. (1988). Complete protein sequences of the variable regions of the cloned heavy and light chains of a human anticytomegalovirus antibody reveal a striking similarity to human monoclonal rheumatoid factors of the Wa idiotypic family. *J. Clin. Invest.* **81**, 1511–1518.

Nickerson, K. G., Berman, J., Glickman, E., Chess, L. & Alt, F. W. (1989). Early human IgH gene assembly in Epstein–Barr virus-transformed fetal B cell lines. Preferential utilization of the most JH-proximal D segment (DQ52) and two unusual VH-related rearrangements. *J. Exp. Med.* **169**, 1391–1403.

Noma, Y., Yaoita, Y., Matsunami, N., Rosen, A., Klein, G. & Honjo, T. (1984). Immunoglobulin gene organization of ultraviolet-illuminated human lymphoblastoid cell lines producing both IgM and IgG. *Mol. Biol. Med.* **2**, 337–350.

Olee, T., Yang, P. M., Siminovitch, K. A., Olsen, N. J., Hillson, J., Wu, J., Kozin, F., Carson, D. A. & Chen, P. P. (1991). Molecular basis of an autoantibody-associated restriction fragment length polymorphism that confers susceptibility to autoimmune diseases. *J. Clin. Invest.* **88**, 193–203.

Pascual, V. & Capra, J. D. (1991). Human immunoglobulin heavy chain variable region genes: organization, polymorphism and expression. *Advan. Immunol.*, vol. 49, pp. 1–147, Academic Press, San Diego.

Pascual, V., Randen, I., Thompson, K., Sioud, M., Forre, O., Natvig, J. & Capra, J. D. (1990). The complete nucleotide sequences of the heavy chain variable regions of six monospecific rheumatoid factors derived from Epstein–Barr virus-transformed B cells isolated from the synovial tissue of patients with rheumatoid arthritis. Further evidence that some autoantibodies are unmutated copies of germ line genes. *J. Clin. Invest.* **86**, 1320–1328.

Pascual, V., Victor, K., Lelsz, D., Spellerberg, M. B., Hamblin, T. J., Thompson, K. M., Randen, I., Natvig, J., Capra, J. D. & Stevenson, F. K. (1991). Nucleotide sequence analysis of the V regions of two IgM cold agglutinins. Evidence that the VH4-21 gene segment is responsible for the major cross-reactive idiotype. *J. Immunol.* **146**, 4385–4391.

Perry, D. J. & Carrell, R. W. (1989). CpG dinucleotides are "hotspots" for mutation in the antithrombin III gene. Twelve variants identified using the polymerase chain reaction. *Mol. Biol. Med.* **6**, 239–243.

Rechavi, G., Bienz, B., Ram, D., Ben, N. Y., Cohen, J. B., Zakut, R. & Givol, D. (1982). Organization and evolution of immunoglobulin VH gene subgroups. *Proc. Nat. Acad. Sci., U.S.A.* **79**, 4405–4409.

Rechavi, G., Ram, D., Glazer, L., Zakut, R. & Givol, D. (1983). Evolutionary aspects of immunoglobulin heavy chain variable region (VH) gene subgroups. *Proc. Nat. Acad. Sci., U.S.A.* **80**, 855–859.

Reynaud, C.-A., Dahan, A., Anquez, V. & Weill, J.-C. (1989). Somatic hyperconversion diversifies the single V_H gene of the chicken with a high incidence in the D region. *Cell*, **59**, 171–183.

Rioux, J. D., Larose, Y., Brodeur, B., Rauch, J. E. & Newkirk, M. M. (1991). Structural relationships between rheumatoid factors and anti-viral antibodies. *American College of Rheumatology 55th Annual Scientic Meeting*, 17–21 November, Boston.

Robbins, D. L., Kenny, T. P., Coloma, M. J., Gavilondo, C. J., Soto, G. R., Chen, P. P. & Larrick, J. W. (1990). Serologic and molecular characterization of a human monoclonal rhematoid factor derived from rheumatoid synovial cells. *Arthritis Rheum.* **33**, 1188–1195.

Roudier, J., Silverman, G. J., Chen, P. P., Carson, D. A. & Kipps, T. J. (1990). Intraclonal diversity in the VH genes expressed by $CD5^-$ chronic lymphocytic leukemia-producing pathologic IgM rheumatoid factor. *J. Immunol.* **144**, 1526–1530.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487–491.

Sam, M., Walter, M. A. & Cox, D. W. (1988). RsaI polymorphism of a human immunoglobulin VH5 subclass locus. *Nucl. Acids Res.* **16**, 8748.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Sanz, I., Casali, P., Thomas, J. W., Notkins, A. L. & Capra, J. D. (1989*a*). Nucleotide sequences of eight human natural autoantibody Vh regions reveals apparent restricted use of VH families. *J. Immunol.* **142**, 4054–4061.

Sanz, I., Dang, H., Takei, M., Talal, N. & Capra, J. D.

(1989b). V_H sequence of a human anti-Sm autoantibody. Evidence that autoantibodies can be unmutated copies of germline genes. *J. Immunol.* **142**, 883–887.

Sanz, I., Kelly, P., Williams, C., Scholl, S., Tucker, P. & Capra, J. D. (1989c). The smaller human VH gene families display remarkably little polymorphism. *EMBO J.* **8**, 3741–3748.

Sasso, E. H., van Dijk, K. W. & Milner, E. C. (1990). Prevalence and polymorphism of human VH3 genes. *J. Immunol.* **145**, 2751–2757.

Sasso, E. H., Silverman, G. J. & Mannik, M. (1991). Human IgA and IgG F(ab')2 that bind to staphylococcal protein A belong to the VHIII subgroup. *J. Immunol.* **147**, 1877–1883.

Schroeder, H. J. & Wang, J. Y. (1990). Preferential utilization of conserved immunoglobulin heavy chain variable gene segments during human fetal life. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 6146–6150.

Schroeder, H. J., Hillson, J. L. & Perlmutter, R. M. (1987). Early restriction of the human antibody repertoire. *Science*, **238**, 791–793.

Schroeder, H. J., Hillson, J. L. & Perlmutter, R. M. (1990). Structure and evolution of mammalian VH families. *Int. Immunol.* **2**, 41–50.

Schutte, M. E., Ebeling, S. B., Akkermans, K. E., Gmelig, M. F. & Logtenberg, T. (1991). Antibody specificity and immunoglobulin VH gene utilization of human monoclonal CD5+ cell lines. *Eur. J. Immunol.* **21**, 1115–1121.

Shen, A., Humphries, C., Tucker, P. & Blattner, F. (1987). Human heavy-chain variable region gene family nonrandomly rearranged in familial chonic lymphocytic leukemia. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 8563–8567.

Shin, E. K., Matsuda, F., Nagaoka, H., Fukita, Y., Imai, T., Yokoyama, K., Soeda, E. & Honjo, T. (1991). Physical map of the 3′ region of the human immunoglobulin heavy chain locus: clustering of autoantibody-related variable segments in one haplotype. *EMBO J.* **10**, 3641–3645.

Shuldiner, A. R., Nirula, A. & Roth, J. (1989). Hybrid DNA artifact from PCR of closely related target sequences. *Nucl. Acids Res.* **17**, 4409.

Silberstein, L. E., Litwin, S. & Carmack, C. E. (1989). Relationship of variable region genes expressed by a human B cell lymphoma secreting pathologic anti-Pr2 erythrocyte autoantibodies. *J. Exp. Med.* **169**, 1631–1643.

Silberstein, L. E., Jefferies, L. C., Goldman, J., Friedman, D., Nowell, P. C., Roelcke, D., Pruzanski, W., Roudier, J. & Silverman, G. J. (1991). Variable region gene analysis of pathologic human autoantibodies to the related i and I red blood cell antigens. *Blood*, **78**, 2372–2386.

Siminovitch, K. A. & Chen, P. P. (1990). The biologic significance of human natural autoimmune responses; relationship to the germline, early immune and malignant B cell variable gene repertoire. *Int. Rev. Immunol.* **5**, 265–277.

Souroujon, M. C., Rubinstein, D. B., Schwartz, R. S. & Barrett, K. J. (1989). Polymorphisms in human H chain V region genes from the VHIII gene family. *J. Immunol.* **143**, 706–711.

Spatz, L. A., Wong, K. K., Williams, M., Desai, R., Golier, J., Berman, J. E., Alt, F. W. & Latov, N. (1990). Cloing and sequence analysis of the VH and VL regions of an anti-myelin/DNA antibody from a patient with peripheral neuropathy and chronic lymphocytic leukemia. *J. Immunol.* **144**, 2821–2828.

Takahashi, N., Noma, T. & Honjo, T. (1984). Rearranged immunoglobulin heavy chain variable region (VH) pseudogene that deletes the second complementarity-determining region. *Proc. Nat. Acad. Sci., U.S.A.* **81**, 5194–5198.

Timmers, E., Kenter, M., Thompson, A., Kraakman, M. E. M., Berman, J. E., Alt, F. W. & Schuurman, R. K. B. (1991). Diversity of immunoglobulin heavy chain gene segment rearrangement in B lymphoblastoid cell lines from X-linked agammaglobulinemia patients. *Eur. J. Immunol.* **21**, 2355–2363.

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature (London)*, **302**, 575–581.

Turnbull, I. F., Bernard, O., Sriprakash, K. S. & Mathews, J. D. (1987). Human immunoglobulin variable region genes: a new VH sequence used to detect polymorphism. *Immunogenetics*, **25**, 184–192.

van der Heijden, R. W. J., Bunschoten, H., Pascual, V., Uytdehaag, F. G. C. M., Osterhaus, A. D. M. E. & Capra, J. D. (1990). Nucleotide sequence of a human monoclonal anti-idiotypic antibody specific for a rabies virus-neutralizing monoclonal idiotypic antibody reveals extensive somatic variability suggestive of an antigen driven immune response. *J. Immunol.* **144**, 2835–2839.

van Dijk, K. W., Sasso, E. H. & Milner, E. C. (1991). Polymorphism of the human Ig VH4 gene family. *J. Immunol.* **146**, 3646–3651.

van Es, J. H., Gmelig Meyling, F. H. J., van de Akker, W. R. M., Aanstoot, H., Derksen, R. H. W. M. & Logtenberg, T. (1991). Somatic mutations in the variable regions of a human IgG anti-double-stranded DNA autoantibody suggest a role for antigen in the induction of systemic lupus erythematosus. *J. Exp. Med.* **173**, 461–470.

Walter, M. A., Surti, U., Hofker, M. H. & Cox, D. W. (1990). The physical organization of the human immunoglobulin heavy chain gene complex. *EMBO J.* **9**, 3303–3313.

White, M. B., Word, C. J., Humphries, C. G., Blattner, F. R. & Tucker, P. W. (1990). Immunoglobulin D switching can occur through homologous recombination in human B cells. *Mol. Cell. Biol.* **10**, 3690–3699.

Winter, G. & Milstein, C. (1991). Man-made antibodies. *Nature (London)*, **349**, 293–299.

Wood, W. I., Gitschier, J., Lasky, L. A. & Lawn, R. M. (1985). Base composition-independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. *Proc. Nat. Acad. Sci., U.S.A.* **82**, 1585–1588.

Wysocki, L. J. & Gefter, M. L. (1989). Gene conversion and the generation of antibody diversity. *Annu. Rev. Biochem.* **58**, 509–531.

Yasui, H., Akahori, Y., Hirano, M., Yamada, K. & Kurosawa, Y. (1989). Class switch from μ to δ is mediated by homologous recombination between $\sigma\mu$ and $\Sigma\mu$ sequences in human immunoglobulin gene loci. *Eur. J. Immunol.* **19**, 1399–1403.

Edited by J. Karn

Note added in proof. Since submission of this paper, we have amplified and cloned six additional V_H segments from DP (DP-75 to DP-80). EMBL Data Library accession numbers for DP-1 to DP-80;

03

Annu. Rev. Med. 1994. 45:491–503

TUMOR NECROSIS FACTOR: A Pleiotropic Cytokine and Therapuetic Target

Kevin J. Tracey, M.D.[1] *and Anthony Cerami, Ph.D.*

Picower Institute for Medical Research, Manhasset, New York 11030

KEY WORDS: shock, cahcexia, immunology

ABSTRACT

Advances in the molecular biology of human diseases indicate that the most striking manifestations of illness may be caused not by exogenous pathogenic or tumor porducts, but rather by toxic peptides produced by the host itself. Tumor necrosis factor (TNF), a polypeptide cytokine produced during infection, injury, or invasion, has proved pivotal in triggering the lethal effects of septic shock syndrome, cachexia, and other systemic manifestations of disease. Because removing TNF from the diseased host may prevent development of the illness, this factor has recently been the focus of intensive research. This review discusses the biology of this cytokine, with particular emphasis on its potential therapeutic role in septic shock and cachexia.

INTRODUCTION

Infection, cancer, or inflammation triggers production of immunological mediators termed cytokines. These hormone-like peptides may enter the bloodstream to alter the biology of distant tissues, or they may behave as paracrine

[1] Laboratory of Biomedical Science, Department of Surgery, Division of Neurosurgery, North Shore University Hospital-Cornell University Medical College, Manhasset, New York 11030

0066-4219/94/0401-0491$05.00

mediators that act only locally. In some diseases, cytokines are benefical to the host, representing an integral component to the immunological defense network. In other cases, however, cytokines actually cause the most striking manifestations of the disease (e.g. shock, tissue injury, weight loss). One cytokine that has been widely studied for its role in human disease is tumor necrosis factor (TNF). TNF produces a broad scope of injurious and benefical effects that make this cytokine especially important as a potential therpeutic target. Because the number of primary research publications in this field is too large to permit comprehensive review, this chapter presents an abridged review of the role of TNF in human disease, with particular emphasis on implications for future therapy.

STRUCTURE, BIOSYNTHESIS, AND RECEPTORS

The term tumor necrosis factor (TNF) refers to two closely related cytokines (encoded by separate genes) known as tumor necrosis factor-α (TNF, cachectin) and tumor necrosis factor-β (lymphotoxin, TNF-β). Both cytokines interact with the same cell membrane receptors, and both have been implicated as pathogenic mediators of human illness. However, because much less is known about the pathobiology of lymphotoxin, the remainder of this discussion is confined to cachectin (TNF).

TNF is synthesized by macrophages and other cells in response to bacterial toxins, inflammatory products, and other invasive stimuli (Table 1). At least three different biological activities have been used to detect TNF in biological samples. These activities include cytotoxicity against tumor cells, suppression of adipocyte lipoprotein lipase (LPL), and reduction of myocyte resting membrane potential (E_m) (Table 1). The gene for human TNF encodes a prohormone that is inserted into the cell membrane as a polypeptide with an apparent molecular weight of 26 kDa (1–3). This membrane-bound form of TNF is bioactive as assayed by cell cytotoxicity and has been implicated in the paracrine activities of TNF in tissues. In response to bacterial endotoxin (lipopolysaccharide, or LPS) and other stimuli, the 26 kDa form is proteolytically cleaved into a 17 kDa polypeptide subunit composed of 157 amino acids (4, 5). Three of these TNF monomers associate noncovalently to form a trimer (6, 7). The resultant bell-shaped trimeric form is the predominate bioactive form of TNF in serum and other bodily fluids.

Two types of TNF receptors, Type I (TNFR-55) and Type II (TNFR-75), are present on the plasma membranes of virtually all cells except the erythrocyte. The receptors share structural homology in the extracellular TNF-binding domains and exhibit similar binding affinity for TNF, but they induce separate cytoplasmic signaling pathways following receptor-ligand binding (8, 9). Receptor-ligand interaction likely leads to the clustering of three receptors to one

Table 1 The biochemistry of TNF

Structure	Membrane-associated form is 26 kD Secreted form is 17 kD Biologically active form is a bell-shaped structure composed of three noncovalently bound 17 kD molecules
Stimuli to TNF release	Bacterial toxins: (lipopolysaccharide, enterotoxin, toxic shock syndrome toxin) Viruses: HIV, influenza Mycobacteria Fungi Parasites Products of complement activation Antigen-antibody complexes Cytokines
Cellular sources	Macrophages Lymphocytes Polymorphonuclear leukocytes Eosinophils Astrocytes Langerhans cells Kupffer cells
Cellular biological activities	Cytotoxic to some tumor cells Growth factor for some tumor cells Suppresses LPL in adipocytes Decreases Em in myocytes
Serum half-life	6–20 minutes in mammals following intravenous bolus injection

TNF trimer (10). This interaction triggers maximal cellular responses when as little as 10% of the cell membrane receptors are occupied. Internalization of the bound ligand is not required for cellular activation (11). A number of postreceptor signaling pathways have been implicated in the cellular mechanisms of TNF-R signal transduction, but the task of assigning individual pathways in the mediation of specific cellular responses is incomplete.

Cleaved fragments of both receptor types, also known as TNF-binding proteins (TNF-BPs), have been detected in the urine and serum of patients with a variety of diseases, including cancer, AIDS, and sepsis (12–16). The biological role of these TNF-BPs is interesting because in some cases they inhibit or neutralize the activity of TNF by preventing receptor-ligand interaction. In other cases, however, the TNF-BPs may prolong or enhance the activity of circulating TNF by stabilizing the trimeric structure of the cytokine, thus preventing clearance from serum (17, 18). Investigators have teleologically reasoned that TNF-BPs are released during periods of acute excessive cytokine production to simultaneously prevent the toxic effects of

hypercytokinemia and to provide a pool of bioactive TNF that is slowly released during the host response to invasion.

BIOLOGICAL EFFECTS OF TNF IN VIVO

The biological and metabolic effects of TNF in vivo have been delineated by investigators employing primarily three experimental approaches: (*a*) direct administration of highly purified recombinant TNF; (*b*) administration of anti-TNF antibodies that neutralize the TNF produced endogenously after LPS, tumors, or other stimuli; and (*c*) administration of TNF with another agent(s) that increases or decreases biological response to TNF. These studies have provided an abundant volume of physiological data about the biology of TNF in experimental animals, volunteer subjects, and patients afflicted with cancer or septic shock syndrome. In general, the biological effects of acute, high-dose systemic TNF exposure differ from the effects of repeated exposure to lower doses of TNF (Table 2). These dose-dependent biological responses, which are discussed separately for clarity of presentation, are not discontinuous but rather represent the spectrum of biological activities triggered by the direct and indirect actions of systemic TNF exposure (reviewed in 19, 20).

Acute Exposure to High Doses of TNF Results in Shock and Tissue Injury

The earliest reported studies of systemic exposure to recombinantly derived human TNF revealed that it causes a syndrome of shock and tissue injury that is virtually indistinguishable from septic shock syndrome (21, 22). Within minutes after acute intravenous or intra-arterial administration of TNF (in quantities that could be produced endogenously by the host during infection), a syndrome of shock, tissue injury, capillary leakage syndrome, hypoxia, pul-

Table 2 Systemic effects of TNF

Acute, high dose	Chronic Low Dose
Shock and tissue injury	Weight loss
Catabolic hormone release	Anorexia
Vascular leakage syndrome	Protein catabolism
Adult respiratory distress syndrome	Lipid depletion
Gastrointestinal necrosis	Hepatosplenomegaly
Acute renal tubular necrosis	Subendocardial inflammation
Adrenal hemorrhage	Insulin resistance
Decreased muscle membrane potentials	Enhanced rate of tumor metastasis
Disseminated intravascular coagulation	Acute phase protein release
Fever	Endothelial activation

monary edema, and multiple organ failure associated with a high mortality ensues. Despite the short circulating serum half-life (6–20 min), the phasic appearance of TNF in the circulation triggers a cascade of classical hormonal and cytokine responses that persist for hours after serum TNF has been cleared and culminate in lethal end organ injury (21, 22). Hemodynamic studies indicate that TNF leads to decreased cardiac output, reduced filling pressures, and diminished ejection fraction as occurs in endotoxic or septic shock (23). Although other cytokines implicated in the development of septic shock syndrome possess some degree of toxicity, TNF is the only known factor sufficient to initiate the complex metabolic, hemodynamic, and pathological sequelae of septic shock syndrome (reviewed in 24).

These responses to TNF are mediated in part by its direct effects and in part by secondary mediators triggered by TNF itself. An abridged list of these secondary mediators includes epinephrine, norepinephrine, glucagon, cortisol, ACTH, growth hormone, interleukins -1 and -6, γ-interferon, platelet-activating factor (PAF), and a number of eicosanoids. Nonsteroidal anti-inflammatory agents that prevent eicosanoid biosynthesis result in an attenuation of the metabolic effects of TNF (e.g. fever, myalgia), but they do not uniformly protect against lethal shock and tissue injury (25, 26). Agents that inhibit PAF may provide protection against TNF-induced gastrointestinal ischemia as well (27). IL-1 receptor antagonists and antibodies against γ-interferon partially abrogate the toxicity of TNF by reducing the synergistic effect of these cytokines on TNF-induced tissue injury (28).

Anti-TNF antibodies that neutralize systemic exposure to TNF prevent the development of shock and tissue injury in animal models of endotoxic and septic shock syndrome (29–34). In addition to preventing the direct effects of TNF itself, anti-TNF antibodies also prevent the development of the TNF-triggered hormonal and cytokine cascade (29, 35). These and other investigations have prompted the evaluation of anti-TNF antibodies in randomized, prospective, multicenter clinical trials in the US and Europe. Although the efficacy results of these trials are unavailable pending completion of the trials, preliminary data indicate that anti-TNF confer a servival advantage to patients with septic shock syndrome.

Chronic TNF Exposure Causes Cachexia and Tolerance

The biological effects of chronic exposure to TNF have been investigated by adminstering TNF via three different schedules: (*a*) repeated injection of TNF; (*b*) continuous infusion of TNF; and (*c*) injection of genetically engineered tumors that secrete TNF constitutively. The net effect of prolonged exposure to TNF by any route is the development of cachexia characterized by anorexia, weight loss, dehydration, and depletion of whole-body protein and lipid. As with the acute biological responses, some chronic effects of TNF are directly

attributable to the cytokine itself, whereas others are the result of secondary factors triggered by TNF.

For example, TNF acts directly in the hypothalamus to mediate fever, anorexia, and the release of corticotropin-releasing factor (CRF) (reviewed in 36). Adipocytes exposed to TNF are depleted of lipid and develop net insulin resistance caused by a suppression of lipogenic enzymes that prevents the incorporation of glucose into lipid (37–39). Myocytes exposed to TNF exhibit a decrease in resting membrane potentials, which contributes to the development of intracellular sodium sequestration and third-space dehydration (22, 40, 41). Other muscle responses to TNF include insulin resistance, lactate release, and net protein degradation (22, 42–44). In contrast, the responses of liver to TNF are more anabolic in nature, with net increases in the biosynthesis of lipids, acute phase proteins, and other cytokines (44, 45).

The magnitude of these TNF-mediated cellular responses depends on the quantity of TNF present in the cellular milieu, but obtaining an accurate measure of TNF levels in tissues has proven exceedingly difficult. For instance, compartmentalized production of TNF in one organ may induce metabolic changes in distant organ systems, thereby confounding an assignment of specific biological effects of TNF. This occurs when large amounts of TNF are produced in the brain and causes profound anorexia without whole-body protein loss. High levels of TNF released from peripheral tissues cause less severe anorexia but more severe protein catabolism (46). Thus the biological effects of TNF are dependent on where it is produced in the body.

During early investigations of the effects of chronic exposure to TNF, we and others discovered that animals receiving TNF by repetitive injection developed a tolerance for its toxic effects (i.e. weight loss and shock) (42, 47, 48). Although the molecular basis for this desensitization phenomenon remains unknown, it cannot be attributed to the formation of anti-TNF antibodies. TNF tachphylaxis may be due to altered expression or function of peripheral TNF receptors or to the release of neutralizing TNF-BPs, but these hypotheses have not been verified. Because TNF is a pivotal mediator of endotoxic shock, it is perhaps not surprising that animals made tolerant to TNF are also protected against subsequent exposure to bacterial endotoxin (LPS) (47). The mechanisms underlying this protected state are believed to result from both decreased tissue reponsiveness to TNF and decreased synthesis of TNF in response to LPS (49–51).

TNF Produced in Tissues Causes Inflammation

TNF is a pivotal mediator of inflammation that activates leukocytes, enhances adherence of neutrophils and monocytes to endothelium, promotes migration of inflammatory cells into the intercellular matrix, stimulates fibroblast proliferation, and triggers local production of other proinflammatory cytokines.

These and numerous other inflammatory effects have been implicated in the beneficial effects of TNF. During the early phases of infection, invasion, neoplastic transformation, injury, or wound healing, these inflammatory actions are thought to contribute to the localization and walling off of infection, cytotoxicity against malignant cells, facilitation of tissue remodeling. Presumably these beneficial effects of the TNF gene, not its toxicity during sudden or chronic overproduction, led to its conservation throughout mammalian evolution.

THE PATHOBIOLOGY OF TNF

Serum TNF Levels

Interpretation of serum TNF levels is hampered by the complex interaction of a number of factors (Table 3). Commonly used TNF assays are based on TNF bioactivity (cytotoxicity against specific cell lines) or TNF immunoreactivity (ELISA). In either case, however, it is often difficult to determine precisely what is being measured. For instance, some TNF-BPs neutralize TNF bioactivity but do not prevent recognition of the bound TNF by ELISA. In other

Table 3 Factors that influence the interpretation of serum TNF levels in health and disease.

Factors	Assessment
Normal range	Unknown; may be elevated in healthy volunteers
Binding proteins	Circulating receptor fragments that bind TNF may neutralize bioactivity in some cases but prolong biological effects in others
Assay variables	ELISAs measure immunoreactivity, but do not indicate bioactivity. Bioassays measure specific activity but do not indicate free TNF levels. Bioassays may not be equivalent to physiological (lethality) assays
Intermittent production	TNF release is discontinuous. Intermittent bolus production occurs throughout the course of protracted illnesses.
Synergy with other mediators	IL-1, γ-interferon, LPS, and other factors increase the biological effects of TNF.
Secondary mediators	TNF triggers the appearance of a cascade of secondary mediators. The biological effects of this cascade are temporally dissociated from TNF itself.
Site of production	Compartmentalized production of TNF (e.g. in brain) mediates systemic effects independent of blood levels.

cases, the TNF-TNF-BP complex may actually exhibit higher bioactivity than the uncomplexed TNF trimer. Another confounding variable in the clinical application of TNF serum assays is that TNF acts in part by triggering a cytokine and hormonal cascade that persists for some time after the intial TNF burst is cleared. These secondary mediators (e.g. IL-1, γ-interferon) may synergistically increase the toxicity of TNF even when TNF is present at levels below the limits of detection of currently available assays. At present there is no available measure of integrated TNF exposure over time that is analagous to hemoglobin-A1c in monitoring chronic hyperglycemia. The development of

Table 4 Abridged List of Diseases in which TNF may participate as a pathobiological mediator.

Disease	Implicated biological effects of TNF
Septic Shock	Hypotension Myocardial suppression Vascular leakage syndrome Organ necrosis Stimulates release of toxic secondary mediators Activates clotting cascades
Cancer	Induces cachexia Stimulates tumor growth Enhances metastatic potential Mediates cytotoxicty
AIDS	Induces cachexia Stimulates viral proliferation Mediates central nervous system injury
Transplantation	Key mediator of allograft rejection and graft vs host disease Triggers veno-occlussive disease
Multiple Sclerosis	Cytotoxicty to oligodendrocytes Induces inflammatory plaques
Diabetes	Promotes death of Islets Mediates insulin resistance
Rheumatoid arthritis	Activation of tissue inflammation Implicated in joint destruction
Trauma	Increases energy expenditure Promotes bacterial translocation Vascular leakage syndrome
Malaria	Brain inflammation Capillary thrombosis and infarction
Meningitis	Brain inflammation Breakdown of blood-brain barrier Triggers septic shock syndrome Activates venous infarction
Adult respiratory distress syndrome	Stimulates leukocyte-endothelial activation Direct cytotoxicity to pneumocytes Vascular leakage syndrome

such a hypothetical test for TNF will be expected to improve systemic TNF production in disease.

TNF IN MEDICINE

Advances in molecular biology and medicine continue to shorten the time between the discovery of a cytokine and the clinical application of this knowledge. Although recombinant TNF has only been readily available to investigators for 10 years, TNF has already been implicated in the pathobiology of dozens of human diseases (Table 4). Moreover, this research has led to international clinical trials addressing the role of TNF in septic shock, cancer, rheumatoid arthritis, malaria, and other afflictions. An abridged discussion of the pathogenic role of TNF in septic shock syndrome and cancer follows.

Septic Shock Syndrome

Whereas early investigators attributed the pathological sequelae of septic shock syndrome to the direct toxicity of the invading organisms, it is now generally believed that most (if not all) of the pathological sequelae of septic shock syndrome are triggered by an acute overproduction of cytokines (52, 53). TNF has been implicated as a pivotal mediator of septic shock syndrome because (*a*) it is acutely overproduced and released systemically during overwhelming sepsis (54); (*b*) administration of TNF itself causes shock and tissue injury indistinguishable from septic shock syndrome (21, 22); (*c*) passive immunization with anti-TNF antibodies prevents septic shock syndrome (29); (*d*) neutralization of TNF in septic shock syndrome attenuates the characteristic cytokine and hormonal cascade (29, 35); and (*e*) transgenic mice with a mutation of TNFR-I are protected against endotoxic shock (55). These and other data suggest that septic shock syndrome is dissociable from the presence of the bacteria themselves because when the biological effects of TNF are neutralized, septic shock syndrome fails to develop despite the presence of live bacteria in the bloodstream (29).

Several strategies to prevent TNF toxicity in septic shock syndrome have been proposed. One approach is to directly inhibit TNF with antibodies or TNF-BPs (16, 29, 33). Another is to inhibit TNF production using glucocorticoids, pentoxifylline, thalidomide, cyclosporine A, or IL-10 (56–59). It has also been suggested that TNF may be removed from the plasma by filtration. Finally, therapy may be targeted against TNF's secondary mediators such as IL-1, γ-interferon, PAF, or the eicosanoids (28, 60–62). At present, a randomized, prospective, multicenter clinical trial for monoclonal anti-TNF antibodies is underway that will hopefully provide data on the efficacy of reducing TNF toxicity in patients with septic shock syndrome.

Cancer

The initial hope that TNF should prove useful as a specific and potent antineoplastic factor has waned. Nonetheless, continuing investigations using high doses of TNF infused directly into a tumor compartment (i.e. liver, extremity, or brain) may yield the desired efficacious result (63). Another experimental approach seeks to identify specific signal transduction pathways unique to TNF cytotoxicity mechanisms that are distinct from its injurious shock-inducing effects. Another strategy is based on the observation that some human tumors produce TNF, which then stimulates tumor growth as an autocrine growth factor. These cases prompted the investigations into the use of anti-TNF strategies to reduce tumor growth in a subset of cancer patients.

TNF has also been implicated in the pathobiology of cancer cachexia, but its role in this metabolic disease is less well defined than in septic shock syndrome. TNF production is increased in macrophages obtained from cancer patients and tumor-bearing animals (64, 65). Prolonged exposure of animals to TNF causes cachexia with typical losses of protein, lipids, and red blood cell mass (42, 46, 66). As in cancer cachexia, TNF-induced cachexia also causes insulin resistance and derangements of glcuose metabolism. Ultimately, the cachectic host dies from loss of body protein, but the molecular basis for TNF-induced muscle protein loss remains unknown. Ongoing investigations address the mechanisms by which TNF causes net protein catabolism in cachexia.

FUTURE IMPLICATIONS

The realization that an endogenous mediator (TNF) is capable of killing the septic patient has catalyzed intensive efforts in drug development. As the mechanisms of TNF-induced toxicity are further delineated, we hope that additional strategies directed toward the molecular basis of shock, tissue injury, and protein loss will emerge.

Literature Cited

1. Kriegler M, Perez C, DeFay K, et al. 1988. A novel form of TNF/cachectin is a cell surface cytotoxic transmembrane protein: ramifications for the complex physiology of TNF. *Cell* 53:45–53
2. Perez C, Albert I, DeFay K, et al. 1990. A nonsecretable cell surface mutant of tumor necrosis factor (TNF) kills by cell-to-cell contact. *Cell* 63:251–58
3. Jue D-M, Sherry B, Luedke C, et al. 1990. Processing of newly synthesized cachectin/tumor necrosis factor in endotoxin-stimu-

lated macrophages. *Biochemistry* 29: 8371–77
4. Pennica D, Nedwin GE, Hayflick JS, et al. 1984. Human tumor necrosis factor: Precursor structure, expression, and homology to lymphotoxin. *Nature* 312:724–29
5. Davis JM, Narachi MA, Alton K, Arakawa T. 1987. Structure of human tumor necrosis factor-α derived from recombinant DNA. *Biochemistry* 26:1322–26
6. Smith RA, Baglioni C. 1987. The active form of tumor necrosis factor is a trimer. *J. Biol. Chem.* 262:6951–54
7. Jones EY, Stuart DI, Walker NP. 1989. Structure of tumour necrosis factor. *Nature* 338:225–28
8. Smith CA, Davis T, Anderson D, et al. 1990. A receptor for tumor necrosis factor defines an unusual family of cellular and viral proteins. *Science* 248:1019–23
9. Thoma B, Grell M, Pfizenmaier K, Scheurich P. 1990. Identification of a 60-kD tumor necrosis factor (TNF) receptor as the major signal transducing component in TNF responses. *J. Exp. Med.* 172:1019–23
10. Banner DW, D'Arcy A, Janes W, et al. 1993. Crystal structure of the soluble human 55 kd TNF receptor-human TNF-β complex: Implications for TNF receptor activation. *Cell* 73:431–45
11. Engelmann H, Holtmann H, Brakebusch C, et al. 1990. Antibodies to a soluble form of a tumor necrosis factor (TNF) receptor have TNF-like activity. *J. Biol. Chem.* 265: 14497–504
12. van Zee KJ, Kohno T, Fischer E, et al. 1992. Tumor necrosis factor soluble receptors circulate during experimental and clinical inflammation and can protect against excessive tumor necrosis factor-α in vitro and in vivo. *Proc. Natl. Acad. Sci. USA* 89: 4845–49
13. Digel W, Porszolt F, Schmid M, et al. 1992. High levels of circulating soluble receptors for tumor necrosis factor in hairy cell leukemia and type B chronic lymphocytic leukemia. *J. Clin. Invest.* 89:1690–93
14. Kalinkovich A, Engelmann H, Harpaz N. 1992. Elevated serum levels of tumor necrosis factor receptors (zTNF-R) in patients with HIV infection. *Clin. Exp. Immunol.* 89:351–55
15. Aderka D, Engelmann H, Hornik V. 1991. Increased serum levels of soluble receptors for tumor necrosis factor in cancer patients. *Cancer Res.* 51:5602–7
16. Ashkenazi A, Marsters SA, Capon DJ. 1991. Protection against endotoxic shock by a tumor necrosis factor receptor immunoadhesin. *Proc. Natl. Acad. Sci. USA* 88:10535–39
17. Bemelmans MHA, Gouma DJ, Buurman WA. 1993. Influence of nephrectomy on tumor necrosis factor clearance in a murine model. *J. Immunol.* 150:2007–17
18. Aderka D, Engelmann H, Maor Y, et al. 1992. Stablization of the bioactivity of tumor necrosis factor by its soluble receptors. *J. Exp. Med.* 175:323–29
19. Tracey KJ. 1992. TNF and other cytokines in the metabolism of septic shock and cachexia. *Clin. Nutr.* 11:1–11
20. Tracey KJ, Vlassara H, Cerami A. 1989. Cachectin/TNF (tumour necrosis factor). *Lancet* 1:1122–26
21. Tracey KJ, Beutler B, Lowry SF, et al. 1986. Shock and tissue injury induced by recombinant human cachectin. *Science* 234:470–74
22. Tracey KJ, Lowry SF, Fahey TJ III, et al. 1987. Cachectin/tumor necrosis factor induces lethal shock and stress hormone responses in the dog. *Surg. Gynecol. Obstet.* 164:415–22
23. Natanson C, Eichenholz PW, Danner RL, et al. 1989. Endotoxin and tumor necrosis factor challenge in dogs simulate the cardiovascular profile of human septic shock. *J. Exp. Med.* 169:823–32
24. Tracey KJ. 1991. Tumor necrosis factor (cachectin) in the biology of septic shock syndrome. *Circ. Shock* 35:123–28
25. Goldberg AL, Kettelhut IC, Furuno K, et al. 1988. Activation of protein breakdown and prostaglandin E^2 production in rat skeletal muscle in fever is signaled by a macrophage product distinct from interleukin 1 or other known monokines. *J. Clin. Invest.* 81:1378–83
26. Kettelhut IC, Fiers W, Goldberg AL. 1987. The toxic effects of tumor necrosis factor *in vivo* and their prevention by cyclooxygenase inhibitors. *Proc. Natl. Acad. Sci. USA* 84:4273–77
27. Hsueh W, Gonzalez-Crussi F, Arroyave JL. 1987. Platelet-activating factor: An endogenous mediator for bowel necrosis in endotoxemia. *FASEB J.* 1:403–5
28. Alexander HR, Doherty GM, Buresh CM, et al. 1991. A recombinant human receptor antagonist to interleukin 1 improves survival after lethal endotoxemia in mice. *J. Exp. Med.* 173:1029–32
29. Tracey KJ, Fong Y, Hesse DG, et al. 1987. Anti-cachectin/TNF monoclonal antibodies prevent septic shock during lethal bacteraemia. *Nature* 330:662–64
30. Mathison JC, Wolfson E, Ulevitch RJ. 1988. Participation of tumor necrosis factor in the mediation of gram-negative bacterial lipopolysaccharide-induced injury in rabbits. *J. Clin. Invest.* 81:1925–37
31. Beutler B, Milsark IW, Cerami AC. 1985. Passive immunization against cachectin/tumor necrosis factor protects mice from lethal effect of endotoxin. *Science* 229: 869–71

32. Emerson TE, Lindsey DC, Jesmok GJ, et al. 1992. Efficacy of monoclonal antibody agsinst TNF-α in an endotoxemic baboon model. *Circ. Shock* 38:75–84
33. Peppel K, Crawford D, Beutler B. 1991. A tumor necrosis factor (TNF) receptor-IgG heavy chain chimeric protein as a bivalent antagonist of TNF activity. *J. Exp. Med.* 174:1483–89
34. Hinshaw L, Olson P, Kuo G. 1989. Efficacy of post-treatment with anti-TNF monoclonal antibody in preventing the pathophysiology and lethality of sepsis in the baboon. *Circ. Shock* 27:362–69
35. Fong YM, Tracey KJ, Moldawer LL, et al. 1989. Antibodies to cachectin/TNF reduce interleukin-1-β and interleukin-6 appearance during lethal bacteremia. *J. Exp. Med.* 170:1627–33
36. Martin SB, Tracey KJ. 1991. Tumor necrosis factor-α (TNF) in neuroimmunology. *Adv. Neuroimmun.* 2:125–38
37. Beutler B, Mahoney J, Le Trang N, et al. 1985. Purification of cachectin, a lipoprotein lipase-suppressing hormone secreted by endotoxin-induced RAW 264.7 cells. *J. Exp. Med.* 161:984–95
38. Torti FM, Dieckmann B, Beutler B, et al. 1985. A macrophage factor inhibits adipocyte gene expression: an in vitro model of cachexia. *Science* 229:867–69
39. Hotamisligil GS, Shargill NS, Spiegelman BM. 1993. Adipose expression of tumor necrosis factor-α direct role in obesity-linked insulin resistance. *Science* 259:87–91
40. Tracey KJ, Lowry SF, Beutler B, et al. 1986. Cachectin/tumor necrosis factor mediates changes of skeletal muscle plasma membrane potential. *J. Exp. Med.* 164: 1368–73
41. Kagan BL, Baldwin RL, Munoz D, Wisnieski BJ. 1992. Formation of ion-permeable channels by tumor necrosis factor-α. *Science* 255:1427–30
42. Tracey KJ, Wei H, Manogue KR, et al. 1988. Cachectin/tumor necrosis factor induces cachexia, anemia, and inflammation. *J. Exp. Med.* 167:1211–27
43. Lang CH, Dobrescu C, Bagby GJ. 1992. Tumor necrosis factor impairs insulin action on peripheral glucose. *Endocrinology* 130:43–52
44. Flores EA, Bistrian BR, Pomposelli JJ, et al. 1989. Infusion of tumor necrosis factor/cachectin promotes muscle catabolism in the rat. A synergistic effect with interleukin 1. *J. Clin. Invest.* 83:1614–22
45. Grunfeld C, Wilking H, Neese R, et al. 1989. Persistence of the hypertriglyceridemic effect of tumor necrosis factor despite development of tachyphylaxis to its anorectic/cachectic effects in rats. *Cancer Res.* 49:2554–60
46. Tracey KJ, Morgello S, Koplin B, et al. 1990. Metabolic effects of cachectin/tumor necrosis factor are modified by site of production: Cachectin/tumor necrosis factor-secreting tumor in skeletal muscle induces chronic cachexia, while implantation in brain induces predominately acute anorexia. *J. Clin. Invest.* 86:2014–24
47. Fraker DL, Stovroff MC, Merino MJ, Norton JA. 1988. Tolerance to tumor necrosis factor in rats and the relationship to endotoxin tolerance and toxicity. *J. Exp. Med.* 168:95–105
48. Socher SH, Friedman A, Martinez D. 1988. Recombinant human tumor necrosis factor induces acute reductions in food intake and body weight in mice. *J. Exp. Med.* 167: 1957–62
49. Cockfield SM, Ramassar V, Halloran PF. 1993. Regulation of IFN-γ and TNF-α expression in vivo. *J. Immunol.* 150:342–52
50. Salkowski CA, Vogel SN. 1992. Lipopolysaccharide inceases glucocorticoid receptor expression in murine macrophages. *J. Immunol.* 149:4041–47
51. Zuckerman SH, Evans GF. 1992. Endotoxin tolerance: in vivo regulation of tumor necrosis factor and interleukin-1 synthesis is at the transcriptional level. *Cell. Immunol.* 140:513–19
52. Tracey KJ, Lowry SF. 1990. The role of cytokine mediators in septic shock. *Adv. Surg.* 23:21–56
53. Tracey KJ, Cerami A. 1993. Tumor necrosis factor (cachectin) in septic/endotoxic shock. *Handb. Mediators Septic Shock*, pp. 291–308
54. Girardin E, Grau GE, Dayer J-M, et al. 1988. Tumor necrosis factor and interleukin-1 in the serum of children with severe infectious purpura. *New Engl. J. Med.* 319: 397–400
55. Pfeffer K, Matsuyama T, Kundig TM, et al. 1993. Mice deificient for the 55 kd tumor necrosis factor receptor are resistant to endotoxin shock, yet succumb to L. monocytogenes infection. *Cell* 73:457–67
56. Giroir BP, Beutler B. 1992. Effect of amrinone on tumor necrosis factor production in endotoxic shock. *Circ. Shock* 36: 200–7
57. Lilly CM, Sandhu JS, Ishizaka A, et al. 1989. Pentoxifylline prevents tumor necrosis factor-induced lung injury. *Am. Rev. Respir. Dis.* 139:1361–68
58. Sampaio EP, Sarno EN, Galilly R, et al. 1991. Thalidomide selectively inhibits TNF-α production by stimulated human moncytes. *J. Exp. Med.* 173:699–703
59. Beutler B, Krochin N, Milsark IW, et al. 1986. Control of cachectin (tumor necrosis factor) synthesis: mechanisms of endotoxin resistance. *Science* 232:977–80
60. McNamara MJ, Norton JA, Nauta RJ, Al-

exander HR. 1993. Interleukin-1 receptor antibody (IL-1rab) protection and treatment against lethal endotoxemia in mice. *J. Surg. Res.* 54:316–21
61. Silva AT, Cohen J. 1992. Role of interferon-γ in experimental gram-negative sepsis. *J. Infect. Dis.* 166:331–35
62. Arend WP, Welgus HG, Thompson RC, Eisenberg SP. 1990. Biological properties of recombinant human monocyte-derived interleukin 1 receptor antagonist. *J. Clin. Invest.* 85:1694–97
63. Baglioni C. 1992. Mechanisms of cytotoxicity, cytolysis,and growth stimulation by TNF. In *Tumor Necrosis Factors: The Molecules and Their Emerging Role in Medicine*, ed. B Beutler, pp. 425–38. New York: Raven
64. Aderka D, Fisher S, Levo Y, et al. 1985. Cachectin/tumour-necrosis-factor production by cancer patients. *Lancet* 2:1190
65. Stovroff MC, Fraker DL, Travis WD, Norton JA. 1989. Altered macrophage activity and tumor necrosis factor: Tumor necrosis and host cachexia. *J. Surg. Res.* 46:462–69
66. Oliff A, Defeo-Jones D, Boyer M, et al. 1987. Tumors secreting human TNF/cachectin induce cachexia in mice. *Cell* 50:555–63

Humanized antibodies

Greg Winter and William J. Harris

Hybridoma technology enabled rodent monoclonal antibodies to be created against human pathogens and cells, but these had limited clinical utility. Protein engineering is now generating antibodies for treatment of infectious disease, autoimmune disease and cancer by 'humanizing' rodent antibodies. Humanized antibodies have improved pharmacokinetics, reduced immunogenicity and have been used to clinical advantage.

The antibody is an adaptor molecule containing binding sites for antigen at one end and for effector molecules at the other, and has evolved to bind to a vast range of antigens. Binding alone may be sufficient to neutralize some toxins and viruses, however, more commonly, the antibody triggers the complement system and cell-mediated killing. Although antibodies are natural therapeutic agents, it has proved difficult to make human monoclonal antibodies (mAbs) by hybridoma technology.

Rodent mAbs unfortunately have serious disadvantages: a short half-life in serum; only some of the different classes can trigger human effector functions; and the mAbs can also elicit an unwanted immune response in patients (human anti-mouse antibodies or HAMA). HAMA can result in enhanced clearance of the antibody from the serum, blocking of its therapeutic effect and hypersensitivity reactions. These problems have prompted the use of protein engineering technologies to 'humanize' rodent mAbs by transplanting antigen-binding sites from rodent to human antibodies. In principle, humanizing allows access to a large pool of well-characterized rodent mAbs for therapy, including those with specificities against human antigens that are difficult to elicit from a human immune response[1].

The engineering of antibodies is facilitated by the modular arrangement of protein domains: the heavy- and light-chain variable (V) domains are responsible for binding to antigen, and the constant domains to effector functions. As both complement and cell-mediated killing require fully glycosylated antibody, the engineered mAbs are expressed in mammalian hosts. Each antibody domain is encoded by a different genetic exon and, to build recombinant antibodies, the exons are pasted together. The exons encoding variable domains (*V* genes) can be cloned from the genomic DNA of a B-cell hybridoma: more conveniently, the *V* genes of hybridomas are isolated from the mRNA by use of the polymerase chain reaction (PCR). The *V* genes are readily linked to those exons encoding constant domains for expression of mAbs[2] (Fig. 1). Expression vectors have been built with both antibody and viral promoters and enhancers, with *V* and *C* genes as different exons ('genomic') or linked together ('cDNA'). Different markers for selection of transformed cells are available and in both myeloma and CHO hosts[3,4] mAb expression is greatly enhanced by amplifying the number of integrated copies, resulting in yields of up to 0.7 g/l in fermenters[5].

Building humanized antibodies

The first generation of humanized antibodies were simple chimaeric mAbs, in which the variable domains of a rodent mAb are transplanted to the constant domains of human antibodies (Fig. 1). This reduces the immunogenicity of the rodent mAb (see below) and allows the effector functions to be selected for the therapeutic application. Thus the human γ1 isotype appears to be the most effective for complement and cell-mediated killing, while the human γ4 isotype appears more suitable for imaging and blocking[6].

The second generation of humanized antibodies were the socalled complementarity determining regions (CDR) – grafted antibodies, in which the antigen-binding loops of the rodent mAb were built into a human antibody. The architecture of each antibody V-domain consists of a β-sheet sandwich surmounted by antigen-binding loops in different antibodies, these CDR loops are hypervariable in sequence (Fig. 2). It is this hypervariability that allows the antibody repertoire to bind a potentially vast array of antigens. By transplanting (or grafting) the CDRs from rodent mAb to human antibody, the antigen-binding site can also be transferred[7]; indeed the same human framework can be used for mounting different antigen-binding sites[7–9]. However, to recreate the antigen-binding site it is also necessary to consider other possible interactions between the β-sheet framework and the loops. With the help of molecular modelling it is possible to design framework substitutions that maintain key contacts with the CDR loops.

For example, with the rat antibody CAMPATH-1 directed against the CDw52 antigen of human lymphocytes, the simple grafting of the CDRs failed to transplant the binding activity to a human antibody. When the three-dimensional folding of the VH-CDR1 loop of the rat antibody and its contacts with the rat framework were modelled by computer graphics, the framework amino acid residue Phe27 was predicted to pack against the loop. However, in the human framework of the CDR-grafted antibody, Phe27 was replaced by Ser27; indeed when Ser27 was mutated to Phe in the CDR-grafted antibody, the binding activity was restored[8]. In other examples, enhancement of antigen affinity was achieved stepwise by combining several

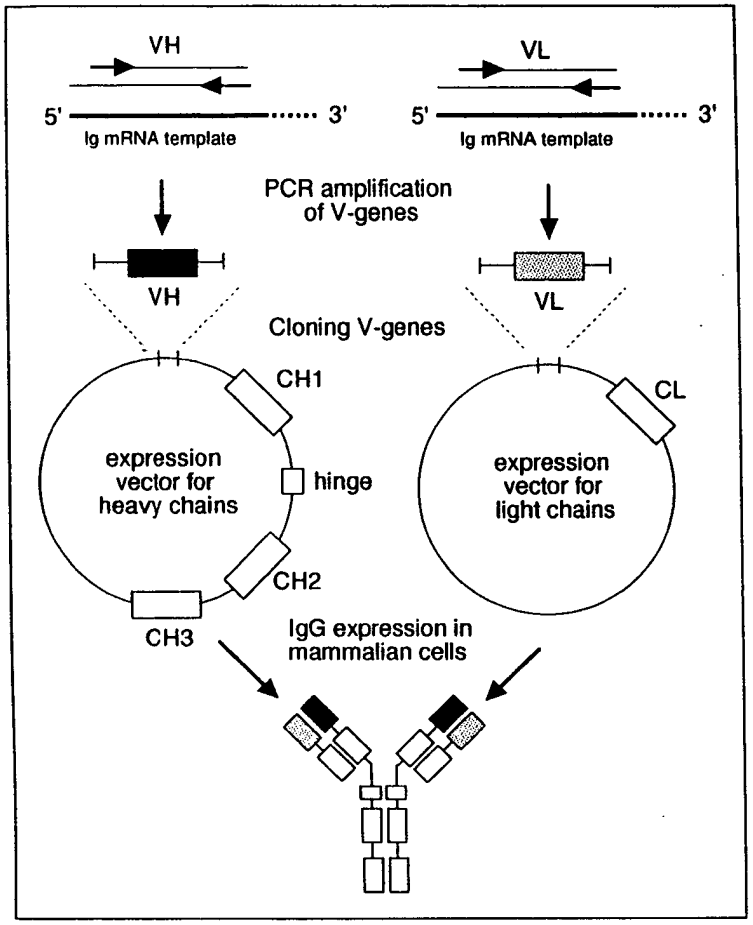


Fig. 1. *Cloning of heavy and light chain V genes from the mRNA of a mouse B-cell hybridoma into vectors comprising (human) genes encoding constant domains, for expression of mouse–human chimaeric antibodies in mammalian cells.*

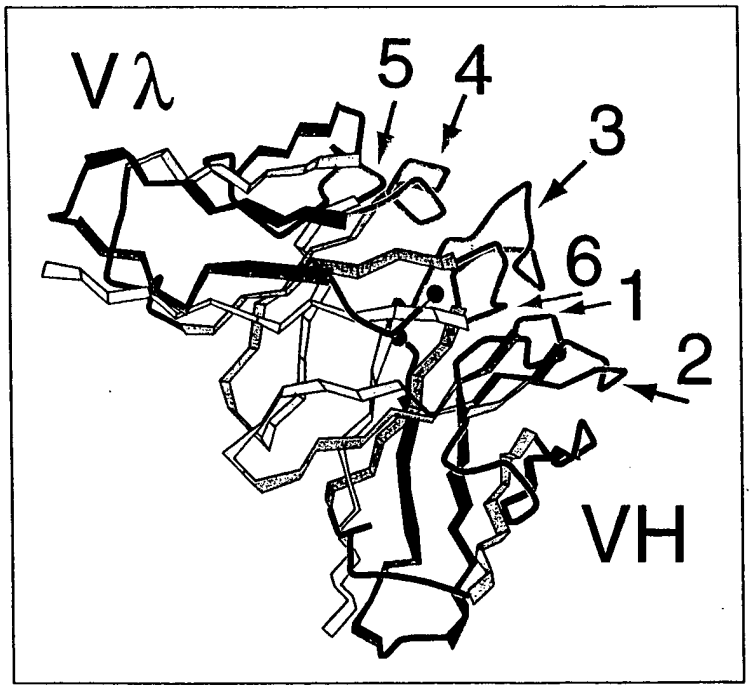


Fig. 2. *The β-sheet framework structure of heavy- and light-chain variable domains with hypervariable loops 1–6. (The figure is taken from Winter, G. and Milstein, C. (1991) and is reproduced with permission*

framework substitutions[10,11]. Indeed analysis of antibody structures is leading towards the identification of sets of framework residues that may exert an influence on CDR structure[11] and also on the packing of the strands of the β-sheet[12].

The first CDR-grafted antibodies were based on the known crystallographic structures of the human myeloma proteins[7–9]. CDR-grafted antibodies have also been built with consensus human frameworks based on several human heavy chains[13]. The use of a single or a limited number of human frameworks offers the prospect of a range of therapeutic antibodies that are almost identical, apart from the CDR sequences. Conversely a range of framework structures should be capable of supporting the CDR loops, and 'hyper-chimaeric' CDR-grafted antibodies have used mouse–human frameworks. For example, to humanize a mouse antibody directed against the human IL-2 receptor (anti-Tac), a human framework sequence was selected by homology. A molecular model was then used to identify those framework residues of the rodent antibody that might interact with the antigen-binding loops, and these were built into the selected human framework[14] (Fig. 3). Chimaeric frameworks have also been proposed in which the internal residues that pack between the domains and with the antigen-binding loops are derived from the rodent sequence and the solvent accessible residues are taken from a human sequence[15].

Most generally, all of the rodent CDRs are transplanted from mouse to human antibody. However, some CDRs are more important than others for binding of antigen, as evident from the crystallographic structures of antibody–antigen complexes. The interaction of antibody loops with antigen involves both main-chain and side-chain contacts: as the CDR loops of mouse and human antibodies fold in a limited number of ways[16], it is possible to maintain some main-chain contacts while varying some of the side chains (and sequence) of the CDRs[17].

There is some loss in binding affinity on CDR-grafting but, in combination with some framework alterations, it is usually possible to obtain a reshaped antibody with an affinity within three fold of the parent monoclonal antibody. High binding affinities may be critical for neutralization of a cytokine or toxin in the serum; they appear to be less important where multiple interactions can occur with high avidities, as with multimeric (cell surface) antibody binding to repeated epitopes on a viral coat[9]. However, small improvements in affinity have been seen for some CDR-grafted antibodies[18], and binding affinities can also be improved *in vitro*, by chain shuffling[19], or random mutation[20].

Using humanized antibodies

Both chimaeric and CDR-grafted antibodies appear to have better pharmacokinetics than rodent mAbs, with extended serum half-life (>75 hours) in humans and cynomolgus monkeys. Likewise the immunogenicity is reduced. Much of the HAMA response to mouse antibodies in patients is directed against the constant region: chimaeric antibodies and CDR-grafted

the immunogenic epitopes being located in the variable regions. Indeed much of this response is directed against the antigen-binding site (for review see Ref. 4).

No antibody response was detected against the CDR-grafted anti-CDw52 antibody during the treatment of two patients with non-Hodgkin's lymphoma for up to 43 days with escalating doses of antibody ranging from 1 to 20 mg/day[21]. Also, no response was reported when the antibody was used in a single course of therapy for rheumatoid arthritis patients[22] or in conjunction with an anti-CD4 antibody in treatment of a patient with intractable systemic vasculitis[23]. However, an antibody response was detected on further treatment of the rheumatoid patients[22]. Antibody responses were not detected against other CDR-grafted antibodies used for radio-imaging in tumour patients[24] or treating acute graft-versus-host disease[25].

Chimaeric and CDR-grafted antibodies have been constructed against a wide range of viral and bacterial pathogens, and against human cell surface markers including tumour cell antigens. Some of the targets are summarized in Table 1. CDR-grafted antibodies against lymphocyte markers have already been used to clinical advantage. The anti-CDw52 antibody was used to deplete a large tumour mass in two lymphoma patients, and to achieve clinical remission[21]. Likewise, this antibody resulted in significant clinical benefit for a patient with systemic vasculitis[23], and for rheumatoid arthritis patients[22]. The anti-Tac antibody was used for immunosuppression following allogeneic marrow transplantation, and resulted in improvement in several cases of acute graft-versus-host disease[25].

Future prospects

As shown above, humanized antibodies can be engineered from rodent mAbs. Their use has been demonstrated in the clinic, and they have a longer serum half-life and reduced immunogenicity compared to rodent mAbs. They therefore appear to be promising as therapeutics, especially for a single course of treatment. It is not yet clear whether humanized (or even human) antibodies will elicit a blocking immune response over longer or several courses of treatment.

The immunogenicity of humanized antibodies is likely to depend on several factors, including the immune state of the patients, and the dose and regime of antibody administration. The target is also likely to be important; antibodies directed against cell-bound antigens might be expected to be more immunogenic than those binding to soluble antigen. Furthermore, since foreign framework regions can elicit an immune response[26], we might also expect that the differing strategies to select and mutate human frameworks, as above, could lead to reshaped antibodies with differing immunogenicity. For example, human antibody frameworks that are mutated (*in vivo* or *in vitro*) with respect to the germ-line segments could prove immunogenic: even buried residues could form the critical element of a T-cell epitope if presented as a denatured peptide by a class II MHC molecule[27]. Indeed it may be desirable to design CDR-grafted antibodies by using framework regions based on human germ-line *V*-gene segments.

Fig. 3. *Computer model of the main-chain backbone of the humanized anti-Tac antibody, with the CDRs in red and altered framework residues in blue (the figure is from prospectus 28 Jan. 1992 of Protein Design Laboratories and reproduced with permission by L. Korn).*

The design of antibodies for therapy would certainly be revolutionised if *in vitro* assays were available to test the immunogenicity of different constructs. However the key practical issues are not whether the immune response can be avoided entirely, but how it

Table 1. CDR-grafted antibodies for therapy (see also Ref. 3)

Target	Clinical potential
CDw52	lymphomas, systemic vasculitis, rheumatoid arthritis
CD3	organ transplantation
CD4	organ transplants, rheumatoid arthritis, Crohn's disease
IL-2 receptor	leukaemias and lymphomas, organ transplants, graft-versus-host disease
TNF-α	septic shock
HIV	AIDS
RSV	respiratory syncytial virus infection
HSV	neonatal, ocular and genital herpes infection
Lewis-Y	cancer
$p185^{HER2}$	cancer
PLAP	cancer
CEA	cancer

TNF: tumour necrosis factor; HIV: human immunodeficiency virus; RSV: rous sarcoma virus; HSV: herpes simplex virus; p185-HER-2: human epidermal growth factor receptor 2; PLAP: placental alkaline phosphatase; CEA: carcinoembryonic antigen.

can be bypassed, for example by changing the idiotype of the engineered antibody, and whether the antibodies can be used for long enough to achieve clinical benefit.

So far we have focussed almost entirely on the construction and use of glycosylated antibodies expressed in mammalian cells. However, the use of antibody fragments may be advantageous in some applications, since they penetrate tissues more readily, and are cleared more rapidly from the serum. This may help in neutralizing and clearing drugs from the serum, or in imaging tumours with radioactive entities coupled to the fragments[1]. Although antibody fragments, lacking the glycosylated Fc portion, cannot trigger effector functions, they could in principle be equipped to do so, for example by chemically linking Fab fragments together[28] as bispecific antibody fragments (with one arm binding a tumour cell antigen and the other binding and triggering effector cells such as cytotoxic T cells or monocytes).

Furthermore, antibody fragments can be expressed by secretion from bacteria[29,30], and can be readily derived from the *V* genes of hybridomas, or from *V* gene repertoires. The repertoires are cloned for display on the surface of filamentous bacteriophage by fusion of the encoded antibody fragment to a coat protein of the phage, and phage with the desired activities selected by binding to antigen. Indeed this technology mimics the strategy of immune selection, and human antibody fragments with specificities against many different foreign and human self-antigens have been isolated from the same 'single pot' of phages (see Refs 31, 32 for review).

Over the past century we have seen three generations of antibody therapeutics: polyclonal animal antibodies, rodent monoclonal antibodies and now humanized antibodies. We anticipate that the use of 'repertoire selection' technologies to make human antibodies and fragments will provide the next generation. However in the meantime it seems likely that humanized antibodies will prove clinically useful for treating several diseases, and the experience should prove valuable for designing and formulating the next generation of antibody therapeutics.

Greg Winter is at the MRC Laboratory of Molecular Biology and the MRC Centre for Protein Engineering, Hills Road, Cambridge, UK CB2 2QH; William Harris is at the Dept of Molecular and Cell Biology, Marischal College, University of Aberdeen, Aberdeen, UK AB9 1AS.

References

1 Winter, G. and Milstein, C. (1991) *Nature* 349, 293–299
2 Orlandi, R., Güssow, D.H., Jones, P.T. and Winter, G. (1989) *Proc. Natl Acad. Sci. USA* 86, 3833–3837
3 Adair, J.R. and Mountain, A. (1992) *Biotech. Gen. Eng. Rev.* 10, 1–142
4 Adair, J.R. (1992) *Immunol. Rev.* 130, 1–40
5 Bebbington, C.R., Renner, G., Thomson, S., King, D., Abrams, D. and Yarranton, G.T. (1992) *Biotechnology* 10, 169–175
6 Morrison, S.L. (1992) *Annu. Rev. Immunol.* 10, 239–265
7 Jones, P.T., Dear, P.H., Foote, J., Neuberger, M.S. and Winter, G. (1986) *Nature* 321, 522–525
8 Riechmann, L., Clark, [illegible] (1988) *Nature* 332, 323–327
9 Tempest, P R. *et al.* (1991) *Biotechnology* 9, 266–271
10 Kettleborough, C.A., Saldanha, J., Heath, V.J., Morrison, C.J., and Bendig, M.M. (1991) *Protein Engin.* 4, 773–783
11 Foote, J. and Winter, G. (1992) *J. Mol. Biol.* 224, 487–499
12 Saul, F.A. and Poljak, R.J. (1993) *J. Mol. Biol.* 230, 15–20
13 Carter, P. *et al.* (1992) *Proc. Natl Acad. Sci. USA* 89, 4285–4289
14 Queen, C. *et al.* (1989) *Proc. Natl Acad. Sci. USA* 86, 10029–10033
15 Padlan, E.A. (1991) *Mol. Immunol.* 28, 489–498
16 Chothia, C. and Lesk, A. (1987) *J. Mol. Biol.* 196, 901–917
17 Glaser, S.M., Vasquez, M., Payne, P.W. and Schneider, W.P. (1992) *J. Immunol.* 149, 2607–2614
18 Co, M.S., Avdalovic, N.M., Caron, P.C. *et al.* (1992) *J. Immunol.* 148, 1149–1154
19 Marks, J.D., Griffiths, A.D., Malmqvist, M. *et al.* (1992) *Biotechnology* 10, 779–789
20 Hawkins, R.E., Russell, S.J. and Winter, G. (1992) *J. Mol. Biol.* 226, 889–896
21 Hale, G. *et al.* (1988) *Lancet* ii, 1394–1399
22 Isaacs, J.D. *et al.* (1992) *Lancet* 340, 748–752
23 Mathieson, P.W. *et al.* (1990) *New Engl. J. Med.* 323, 250–254
24 Hird, V. *et al.* (1991) *Br. J. Cancer* 64, 911–914
25 Anasetti, C. *et al.* (Abstr.) American Society of Haematology, Anaheim, November 1992
26 Brüggemann, M., Winter, G., Waldmann, H. and Neuberger, M.S. (1989) *J. Exp. Med.* 170, 2153–2157
27 Allen, P.M., Masueda, G.R., Haber, E. and Unanue, E.R. (1985) *J. Immunol.* 135, 368–373
28 Carter, P. *et al.* (1992) *Biotechnology* 10, 163–167
29 Skerra, A. and Pluckthun, A. (1988) *Science* 240, 1038–1041
30 Better, M., Chang, C.P., Robinson, R.R. and Horwitz, A.G. (1988) *Science* 240, 1041–1043
31 Marks, J.D., Hoogenboom, H.R., Griffiths, A.D. and Winter, G. (1992) *J. Biol. Chem.* 267, 1–4
32 Hoogenboom, H.R., Marks, J.D., Griffiths, A.D. and Winter, G. (1992) *Immunol. Rev.* 13), 41–68

Letters to the Editor

Immunology Today welcomes your correspondence. Letters to the Editor (maximum 400 words) may address issues raised in *Immunology Today* or other matters of interest to immunologists.

All correspondence should be addressed to The Editor, *Immunology Today*, Elsevier Trends Journals, 68 Hills Road, Cambridge, UK CB2 1LA.

The decision to publish rests with the editor and the author of any *IT* article criticized in a letter will normally be invited to reply.